

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 14, 2006, 17:03:40 ; Search time 41.8457 Seconds

(without alignments)
1133.996 Million cell updates/sec

Title: US-10-620-246-4

Perfect score: 538

Sequence: 1 MADPSTALPDAGLVVL.....ADLADGVALGDGVRSIEKT 108

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_21.*
1: geneseqp1980a:*
2: geneseqp1990a:*
3: geneseqp2000a:*
4: geneseqp2001a:*
5: geneseqp2002a:*
6: geneseqp2003a:*
7: geneseqp2003ba:*
8: geneseqp2004a:*
9: geneseqp2005a:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	534	99.3	1041	6	ABU34693
2	534	99.3	1041	6	ABU36647
3	533	99.1	107	2	AAW72885
4	531	98.7	108	2	AAW72885
5	485	90.1	1045	2	AAW57448
6	449	83.5	1061	6	ABU34060
7	423	78.6	1059	6	ABU35902
8	266.5	49.5	1052	6	ABU26087
9	266.5	49.5	1054	6	AAAG92098
10	266.5	49.5	1063	7	ADDD13268
11	185	34.4	1103	5	ABP66334
12	132	24.5	1091	6	ABU48588
13	131.5	24.4	1065	8	ADN46385
14	128.5	23.9	1035	6	ABU24074
15	127.5	23.7	1185	4	AAU43003
16	127.5	23.7	1185	4	ABM39522
17	123.5	23.0	1067	4	AAAB96257
18	112.5	20.9	1080	7	ADM26293
19	105	19.5	1040	2	AAW97216
20	105	19.5	1041	7	AAV56782
21	105	19.5	1041	7	ADC34773
22	104	19.3	539	2	AAV37205
23	99	18.4	1088	2	AAV00891
24	99	18.4	1088	2	AAV00892

25	98	18.2	1088	5	ABP73361	ABP73361	Candida a
26	97.5	18.1	1036	6	ABU27203	ABU27203	Protein e
27	97	18.0	1026	8	ADSO8184	ADSO8184	Staphyloc
28	97	18.0	1085	6	ABU20747	ABU20747	Protein e
29	96	17.8	1072	6	ABR52806	ABR52806	Protein s
30	96	17.8	1072	7	ADK62202	ADK62202	Disease t
31	93	17.3	1042	6	ABU19362	ABU19362	Protein e
32	89	16.5	1363	6	ABU08100	ABU08100	Human Kin
33	87	16.2	1038	6	ABU25711	ABU25711	Protein e
34	86	16.0	1232	9	ABM95065	ABM95065	M. xanthu
35	82.5	15.3	659	8	ADG29586	ADG29586	Bacterial
36	82.5	15.3	7068	4	AAK10142	AAK10142	Streptomy
37	82.5	15.3	9477	4	AAE10144	AAE10144	Streptomy
38	82.5	15.3	11096	4	AAE10129	AAE10129	Streptomy
39	80.5	15.0	348	5	ABP26034	ABP26034	Streptoco
40	80.5	15.0	348	8	ADV89673	ADV89673	Streptoco
41	80.5	15.0	348	8	ADV83074	ADV83074	Streptoco
42	80.5	15.0	348	8	ADV80926	ADV80926	Streptoco
43	79	14.7	375	8	AD137312	AD137312	M. tuberc
44	78.5	14.6	329	6	ABU34127	ABU34127	Protein e
45	78.5	14.6	1077	6	ABJ26121	ABJ26121	Aspergill

ALIGNMENTS

RESULT 1	ABU34693	ABU34693 standard; protein; 1041 AA.
ID	ABU34693	
AC	ABU34693	
XX		
DT	19-JUN-2003	(first entry)
XX		
DE	Protein encoded by Prokaryotic essential gene #20220.	
KW	Antisense; prokaryotic essential gene; cell proliferation; drug design.	
XX		
OS	Mycobacterium bovis.	
XX		
PN	MO200277183-A2.	
XX		
PD	03-OCT-2002.	
XX		
PF	21-MAR-2002; 2002MO-US009107.	
XX		
PR	21-MAR-2001; 2001US-00815242.	
PR	06-SEP-2001; 2001US-00948993.	
PR	25-OCT-2001; 2001US-0342923P.	
PR	08-FEB-2002; 2002US-00072851.	
PR	06-MAR-2002; 2002US-0362699P.	
XX		
PA	(BLIT-) ELITRA PHARM INC.	
XX		
PI	Wang L, Zamudio C, Malone C, Haasebeck R, Ohlsen KU, Zyskind JW,	
PI	Wall D, Trawick JD, Carr GJ, Yamamoto R, Foreyth RA, Xu HH;	
XX	WPI; 2003-029926/02.	
DR	N-PSDB; ACA38563.	
XX		
PT	New antisense nucleic acids, useful for identifying proteins or screening	
PT	for homologous nucleic acids required for cellular proliferation to	
PT	isolate candidate molecules for rational drug discovery programs.	
XX		
PS	Claim 25; SEQ ID NO 62617; 1766pp; English.	
XX		
CC	The invention relates to an isolated nucleic acid comprising any one of	
CC	the 6213 antisense sequences given in the specification where expression	
CC	of the nucleic acid inhibits proliferation of a cell. Also included are:	
CC	(1) a vector comprising a promoter operably linked to the nucleic acid	
CC	encoding a polypeptide whose expression is inhibited by the antisense	
CC	nucleic acid; (2) a host cell containing the vector; (3) an isolated	
CC	polypeptide or its fragment whose expression is inhibited by the	

CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 1041 AA;

Query Match 99.3%; Score 534; DB 6; Length 1041;
 Best Local Similarity 99.1%; Pred. No. 1.7e-54;
 Matches 107; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAADPESTALPDGAGLVVDGVTVALEBAEGWAKDRIEQLRKSTGLDVSDDRIRVVM 60
 :|||||
 Db 934 VAADPESTALPDGAGLVVDGVTVALEBAEGWAKDRIEQLRKSTGLDVSDDRIRVVM 993
 QY 61 SVPAEREDWARTRDILAGEIILATDFEPADLADGVAIGGVRSIEKT 108
 :|||||
 Db 994 SVPAEREDWARTRDILAGEIILATDFEPADLADGVAIGGVRSIEKT 1041

RESULT 2

ABU36647
 ID ABU36647 standard; protein; 1041 AA.

XX AC ABU36647;

XX DT 19-JUN-2003 (first entry)

DE Protein encoded by Prokaryotic essential gene #22174.

XX KM Antisense; prokaryotic essential gene; cell proliferation; drug design.

XX OS Mycobacterium tuberculosis.

XX PN W0200277183-A2.

XX PD 03-OCT-2002.

XX PF 21-MAR-2002; 2002MO-US009107.

XX PR 21-MAR-2001; 2001US-00815242.

XX PR 06-SEP-2001; 2001US-00948993.

XX PR 25-OCT-2001; 2001US-0342923P.

XX PR 08-FEB-2002; 2002US-00072851.

XX PR 06-MAR-2002; 2002US-0362699P.

XX PA (ELIT-) ELITRA PHARM INC.

XX PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;

XX PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Foreyth RA, Xu HH;

XX DR WPI, 2003-029926/02.

XX DR N-PSDB; ACA40517.

XX PT New antisense nucleic acids, useful for identifying proteins or screening
 CC for homologous nucleic acids required for cellular proliferation to
 CC isolate candidate molecules for rational drug discovery programs.

XX PS Claim 25; SEQ ID NO 64571; 1766bp; English.

XX CC The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 1041 AA;

Query Match 99.3%; Score 534; DB 6; Length 1041;
 Best Local Similarity 99.1%; Pred. No. 1.7e-54;
 Matches 107; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAADPESTALPDGAGLVVDGVTVALEBAEGWAKDRIEQLRKSTGLDVSDDRIRVVM 60
 :|||||
 Db 934 VAADPESTALPDGAGLVVDGVTVALEBAEGWAKDRIEQLRKSTGLDVSDDRIRVVM 993
 QY 61 SVPAEREDWARTRDILAGEIILATDFEPADLADGVAIGGVRSIEKT 108
 :|||||
 Db 994 SVPAEREDWARTRDILAGEIILATDFEPADLADGVAIGGVRSIEKT 1041

RESULT 3

AAW72885
 ID AAW72885 standard; protein; 107 AA.

XX AC AAW72885;

XX DT 21-JUN-1999 (first entry)

DE Mycobacterium tuberculosis antigen CFP9.

XX KM Mycobacterium tuberculosis; antigen; vaccine; immunological; immunogen;

XX KM infection.

XX OS Mycobacterium tuberculosis.

XX PN W09844119-A1.

XX PD 08-OCT-1998.


```

PF 01-APR-1998; 98WO-DK000132.
XX
XX 02-APR-1997; 97DK-00000376.
PR 18-APR-1997; 97US-0044624P.
PR 10-NOV-1997; 97DK-00001277.
PR 05-JAN-1998; 98US-0070488P.
XX
XX (STAT-) STATENS SERUM INST.
XX
XX Andersen P, Nielsen R, Rosenkrands I, Meldingh K, Rasmussen PB,
PI Oettinger T, Florio W;
XX
XX WPI; 1998-542705/46.
DR N-PSDB; AAV63916.
XX
XX New isolated mycobacteria polypeptides and nucleic acids - used for
PT developing products for the diagnosis of or vaccination against
PT mycobacterial infections, particularly tuberculosis.
XX
XX Claim 1; Page 125; 163pp; English.
XX
XX The present sequence represents a Mycobacterium tuberculosis protein.
CC Products from the present invention, which describes protein fragments
CC and nucleic acid fragments derived from M. tuberculosis, can be used in
CC the detection of and prevention of mycobacterial infections. In
CC particular, the proteins and nucleic acids can be used for the diagnosis
CC of or vaccination against tuberculosis caused by M. tuberculosis, M.
CC africanum or M. bovis
XX
XX Sequence 107 AA;
SQ
Query Match 99.1%; Score 533; DB 2; Length 107;
Best Local Similarity 100.0%; Pred. No. 1.1e-55;
Matches 107; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 AADPESTALPDGAGLVLDGTVTALBAGNAKRIRELQELKSTGLDVSRIRYVM 61
DB 1 AADPESTALPDGAGLVLDGTVTALBAGNAKRIRELQELKSTGLDVSRIRYVM 60
QY 62 VPAEREDMARTHRDLIAGEIILATDFEPADLADGVAIGDGVRSIEKT 108
DB 61 VPAEREDMARTHRDLIAGEIILATDFEPADLADGVAIGDGVRSIEKT 107
RESULT 4
AAV21902
ID AAV21902 standard; protein; 108 AA.
XX
XX AAV21902;
AC
XX
XX 06-SEP-1999 (first entry)
DT
XX
XX Amino acid sequence of antigen CPP9.
DE
XX
XX Immunogenic; Mycobacterium tuberculosis; immune response; infection;
KM tuberculosis; fusion polypeptide; T-cell epitope; ESAT-6; MPT59; TB;
KM pharmaceutical; vaccination; M. africanum; M. bovis; CPP7A; CPP30A;
KM CPP7B; CPP19; CPP27; CPP30A; RD1-ORF; CPP10A; CPP16; CPP19; CPP23;
KM CPP25A; CPP30B; CPP7B.
XX
XX Mycobacterium tuberculosis.
OS
XX
XX WO924577-A1.
PN
XX
XX 20-MAY-1999.
PD
XX
XX 08-OCT-1998; 98WO-DK000438.
PF
XX
XX 10-NOV-1997; 97DK-00001277.
PR
XX
XX 05-JAN-1998; 98US-0070488P.
PR
XX
XX 01-APR-1998; 98WO-DK000132.
XX
XX (STAT-) STATENS SERUM INST.

```

```

XX
XX Andersen P, Skjot R;
PI
XX
XX WPI; 1999-347282/29.
DR
XX
XX N-PSDB; AAX61002.
XX
XX New immunogenic fragment of Mycobacterium tuberculosis.
PT
XX
XX Example 1; Fig 4; 265pp; English.
PS
XX
XX The invention describes a substantially pure immunogenic polypeptide
CC fragment (I) from Mycobacterium tuberculosis that is able to evoke a
CC protective immune response against infections by mycobacteria belonging
CC to the tuberculosis complex. The invention provides a (1) fusion
CC polypeptide comprising at least one polypeptide fragment (i) and at least
CC one fusion partner; (2) a fusion polypeptide fragment comprising a T-cell
CC epitope from M. tuberculosis protein ESAT-6, or MPT59 and a second
CC different amino acid sequence from M. tuberculosis, and/or including a
CC sequence which protects the first amino acid sequence from in vivo
CC degradation or post-translational processing; (3) a nucleic acid fragment
CC that encodes the above polypeptides. The polypeptides and nucleic acid
CC are useful as pharmaceuticals, for diagnosis of and as antigens for
CC vaccination against TB caused by Mycobacterium tuberculosis; africanum or
CC bovis. The polypeptides are also useful for diagnosing ongoing or
CC previous sensitization in an animal with bacteria belonging to the
CC tuberculosis complex. The invention also describes the use of CPP7A or
CC CPP30A or a T-cell epitope of for the induction of a strong immune
CC response in a mammal; use of CPP7B, CPP19 or MPT59-ESAT6 or a T-cell
CC epitope of for diagnosis of TB in a mammal by performing a DTH type skin
CC test; use of CPP27, CPP30A, RD1-ORF2, RD1-ORF3, RD1-ORF5, MPT59-ESAT6,
CC ESAT6-MPT59, CPP10A, CPP16, CPP19, CPP23, CPP25A, CPP30B, CPP7B or a T-
CC cell epitope of for the preparation of an immunological composition; and
CC for the preparation of a subunit vaccine
XX
XX Sequence 108 AA;
SQ
Query Match 98.7%; Score 531; DB 2; Length 108;
Best Local Similarity 99.1%; Pred. No. 1.9e-55;
Matches 107; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MAADPESTALPDGAGLVLDGTVTALBAGNAKRIRELQELKSTGLDVSRIRYVM 60
DB 1 MAADPESTALPDGAGLVLDGTVTALBAGNAKRIRELQELKSTGLDVSRIRYVM 60
QY 61 SVPAREDMARTHRDLIAGEIILATDFEPADLADGVAIGDGVRSIEKT 108
DB 61 SVPAREDMARTHRDLIAGEIILATDFEPADLADGVAIGDGVRSIEKT 108
RESULT 5
AAW57448
ID AAW57448 standard; protein; 1045 AA.
XX
XX AAW57448;
AC
XX
XX 09-SEP-2004 (revised)
DT
XX
XX 12-AUG-1998 (first entry)
DT
XX
XX M. tuberculosis isoleucyl-tRNA synthetase (ILER5) enzyme.
DE
XX
XX Isoleucyl-tRNA synthetase; Iler5s; Mycobacterium tuberculosis;
KM antimicrobial; tester strain; leprosy; tuberculosis.
KM
XX
XX Mycobacterium tuberculosis.
OS
XX
XX Unidentified.
OS
XX
XX Key Location/Qualifiers
FH
XX
XX MISC-difference 11
FT
XX
XX /note="If the first GTG codon at position 640 is used
FT for initiation of translation, the GTG codon at position
FT 670 would be translated as Valine at this position to
FT yield a 1045 amino acid protein; if the GTG codon at
FT position 670 is used for initiation, it would be

```


Db 1014 SVPAERADWVHDFIATGRIATPFRGEPADVAIGDGVRSISK 1060

RESULT 7

ABU35902 ID ABU35902 standard; protein; 1059 AA.

XX ABU35902;

DT 19-JUN-2003 (first entry)

DE Protein encoded by Prokaryotic essential gene #21429.

XX Antisense; prokaryotic essential gene; cell proliferation; drug design.

OS Mycobacterium leprae.

XX NO200277183-A2.

PN 03-OCT-2002.

PF 21-MAR-2002; 2002WO-US009107.

PR 21-MAR-2001; 2001US-00815242.

PR 06-SEP-2001; 2001US-00948993.

PR 25-OCT-2001; 2001US-0342923P.

PR 08-FEB-2002; 2002US-00072851.

PR 06-MAR-2002; 2002US-0362699P.

XX (BLIT-) ELITRA PHARM INC.

XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KU, Zyskind JW;

PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Foreyth RA, Xu HH;

XX WPI; 2003-029926/02.

DR N-PSDB; ACA39772.

XX New antisense nucleic acids, useful for identifying proteins or screening

PT for homologous nucleic acids required for cellular proliferation to

PT isolate candidate molecules for rational drug discovery programs.

XX Claim 25; SEQ ID NO 63826; 1766pp; English.

XX The invention relates to an isolated nucleic acid comprising any one of

CC the 6213 antisense sequences given in the specification where expression

CC of the nucleic acid inhibits proliferation of a cell. Also included are:

CC (1) a vector comprising a promoter operably linked to the nucleic acid

CC encoding a polypeptide whose expression is inhibited by the antisense

CC nucleic acid; (2) a host cell containing the vector; (3) an isolated

CC polypeptide or its fragment whose expression is inhibited by the

CC antisense nucleic acid; (4) an antibody capable of specifically binding

CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular

CC proliferation or the activity of a gene in an operon required for

CC proliferation; (7) identifying a compound that influences the activity of

CC the gene product or that has an activity against a biological pathway

CC required for proliferation, or that inhibits cellular proliferation; (8)

CC identifying a gene required for cellular proliferation or the biological

CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 1059 AA;

XX Query Match 78.6%; Score 423; DB 6; Length 1059;

XX Best Local Similarity 80.4%; Pred. No. 3.7e-41;

XX Matches 86; Conservative 6; Mismatches 15; Indels 0; Gaps 0;

QY 1 MAADPESTAAIPDAGLVLDGVTYAELEAGAKRIRIRLOLRKSTGLDVSRRIRVM 60

Db 952 VAANPEFTAEIPDGGVLVDPTVPELEAGAKRIRIRLOLRKSTGLDVSRRIRVM 1011

QY 61 SVPAERADWVHDFIATGRIATPFRGEPADVAIGDGVRSISK 107

Db 1012 SVPAERADWVHDFIATGRIATPFRGEPADVAIGDGVRSISK 1058

RESULT 8

ABU26087 ID ABU26087 standard; protein; 1052 AA.

XX ABU26087;

DT 19-JUN-2003 (first entry)

DE Protein encoded by Prokaryotic essential gene #11614.

XX Antisense; prokaryotic essential gene; cell proliferation; drug design.

OS Corynebacterium diphtheriae.

XX NO200277183-A2.

PN 03-OCT-2002.

PF 21-MAR-2002; 2002WO-US009107.

PR 21-MAR-2001; 2001US-00815242.

PR 06-SEP-2001; 2001US-00948993.

PR 25-OCT-2001; 2001US-0342923P.

PR 08-FEB-2002; 2002US-00072851.

PR 06-MAR-2002; 2002US-0362699P.

XX (BLIT-) ELITRA PHARM INC.

XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KU, Zyskind JW;

PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Foreyth RA, Xu HH;

XX WPI; 2003-029926/02.

DR N-PSDB; ACA29957.

XX New antisense nucleic acids, useful for identifying proteins or screening

PT for homologous nucleic acids required for cellular proliferation to

PT isolate candidate molecules for rational drug discovery programs.

XX Claim 25; SEQ ID NO 54011; 1766pp; English.

XX The invention relates to an isolated nucleic acid comprising any one of

CC the 6213 antisense sequences given in the specification where expression

CC of the nucleic acid inhibits proliferation of a cell. Also included are:

CC (1) a vector comprising a promoter operably linked to the nucleic acid

CC encoding a polypeptide whose expression is inhibited by the antisense

CC nucleic acid; (2) a host cell containing the vector; (3) an isolated

CC polypeptide or its fragment whose expression is inhibited by the

CC antisense nucleic acid; (4) an antibody capable of specifically binding

CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular

CC proliferation or the activity of a gene in an operon required for

CC proliferation; (7) identifying a compound that influences the activity of

CC the gene product or that has an activity against a biological pathway

CC required for proliferation, or that inhibits cellular proliferation; (8)

CC identifying a gene required for cellular proliferation or the biological

CC pathway in which a proliferation-required gene or its gene product lies

CC or a gene on which the test compound that inhibits proliferation of an

CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

XX
 CC Sequence 1052 AA;

Query Match 49.5%; Score 266.5; DB 6; Length 1052;
 Best Local Similarity 49.6%; Pred. No. 2.2e-22;
 Matches 57; Conservative 18; Mismatches 25; Indels 15; Gaps 2;

QY 1 MADPESTALPGAGLVLDGVTYATLEBEGAKRIRRELQKSTGIDVSDRIRYVM 60
 DB 944 VAADPSTQIDVDGLVLDMEVTELEBEGMAADVIRGLQARKASGFVSDRIEVLK 1003
 QY 61 SVPAEREDMARTRDILAGEIILATDPEFADLADGVAIGD-----GVASVIEK 107
 DB 1004 VVEBEKKEMALRTHTDMLAGEVLATSPF-----VTGGEPAHDIYAGVATATQK 1051

RESULT 9
 AAG92098
 ID AAG92098 strand; protein; 1054 AA.

XX AAG92098;
 XX
 XX 26-SEP-2001 (first entry)
 XX
 XX C glutamicum protein fragment SEQ ID NO: 5852.
 XX
 XX Coryneform bacterium; amino acid synthesis; vitamin; saccharide;
 XX organic acid synthesis.
 XX
 XX Corynebacterium glutamicum.
 XX
 XX EPI108790-A2.
 XX
 XX 20-JUN-2001.
 XX
 XX 18-DEC-2000; 2000EP-00127688.
 XX
 XX 16-DEC-1999; 99UP-00377484.
 XX
 XX 07-APR-2000; 2000UP-00159162.
 XX
 XX 03-AUG-2000; 2000JP-00280988.
 XX
 XX (KYOW) KYOWA HAKKO KOGYO KK.
 XX
 XX Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
 XX Tateishi N, Senoh A, Ikeda M, Ozaki A;
 XX
 XX WPI; 2001-376931/40.
 XX
 XX N-PSDB; AAH67317.
 XX
 XX Novel polynucleotides derived from Coryneform bacteria, for identifying
 XX mutation point of a gene, measuring expression of a gene, analyzing
 XX expression profile or pattern of a gene and identifying homologous gene.
 XX
 XX Claim 17, SEQ ID NO 5852; 246bp + Sequence Listing; English.
 XX
 XX The present invention provides a number of nucleotide and protein
 XX sequences from the Coryneform bacterium Corynebacterium glutamicum. These
 XX are useful for identifying the mutation point of a gene derived from a

CC mutant of coryneform bacterium, measuring expression amount and analysing
 CC the expression profile or expression pattern of a gene derived from
 CC Coryneform bacterium, and identifying a homologue of a gene derived from
 CC coryneform bacterium. Coryneform bacteria are useful for producing amino
 CC acids, nucleic acids, vitamins, saccharides and organic acids,
 CC particularly L-lysine. The present sequence is a protein described in the
 CC exemplification of the invention. Note: The sequence data for this patent
 CC did not form part of the printed specification, but was obtained in
 CC electronic format directly from the European Patent Office

XX
 CC Sequence 1054 AA;

Query Match 49.5%; Score 266.5; DB 4; Length 1054;
 Best Local Similarity 53.6%; Pred. No. 2.2e-22;
 Matches 59; Conservative 15; Mismatches 33; Indels 3; Gaps 2;

QY 1 MADPESTALPGAGLVLDGVTYATLEBEGAKRIRRELQKSTGIDVSDRIRYVM 60
 DB 944 VAANPSTQIDVDGLVLDMEVTELEBEGMAADVIRGLQARKASGFVSDRIEVLK 1003
 QY 61 SVPAEREDMARTRDILAGEIILATDPEFADLADGVA--IGDGVRSVIEK 107
 DB 1004 SVPEDEKKEMITTHADHIAAEVLATSPFVIYTDALDGETTHIIVAGVATATK 1053

RESULT 10
 ADD13268
 ID ADD13268 strand; protein; 1063 AA.
 XX
 XX ADD13268;
 XX
 XX 01-JAN-2004 (first entry)
 XX
 XX C. glutamicum stability and folding associated protein RXA02726.
 XX
 XX genetic stability; DNA repair; recombination; transposition;
 XX gene expression; protein folding; fine chemical production;
 XX lysine production; nucleotide production; nucleoside production;
 XX lipid production; diol production; carbohydrate production;
 XX aromatic compound production; vitamin production; co-factor production;
 XX enzyme production; food; animal feed; cosmetic; pharmaceutical.
 XX
 XX Corynebacterium glutamicum.
 XX
 XX
 XX Key Location/Qualifiers
 XX FT Misc-difference 286
 XX FT /note="Optionally substituted with Val"
 XX
 XX WO2003040180-A2.
 XX
 XX 15-MAY-2003.
 XX
 XX 31-OCT-2002; 2002WO-BP012138.
 XX
 XX 05-NOV-2001; 2001DE-01054180.
 XX
 XX (BADI) BASF AG.
 XX
 XX Zelder O, Pompejus M, Schroeder H, Kroege B, Kloppeggge C;
 XX Habermann G;
 XX
 XX WPI; 2003-505062/47.
 XX
 XX N-PSDB; ADD13267.
 XX
 XX New nucleic acid encoding variant forms of proteins required for e.g.
 XX genetic stability and proper protein folding, useful for production of
 XX fine chemicals, specifically lysine, in microorganisms.
 XX
 XX Claim 1, SEQ ID NO 76; 265bp; German.
 XX
 XX This invention describes novel polynucleotides and polypeptides involved
 XX in genetic stability (DNA repair and recombination, transposition of
 XX genetic material), gene expression and folding of proteins in

CC Corynebacterium glutamicum. Polynucleotides are isolated from a nucleic acid library of C. glutamicum then mutated at the specified positions, CC cloned and expressed by standard methods. Cells containing vectors that CC express the polynucleotides are used for production of fine chemicals, CC preferably amino acids and specifically lysine, but more generally CC nucleotides, nucleosides, lipids, fatty acids, diols, carbohydrates, CC aromatic compounds, vitamins, co-factors and enzymes. These are useful in CC the food, animal feed, cosmetics and pharmaceutical industries. The CC polynucleotides, optionally as primers and probes, can also be used for CC identification and classification of C. glutamicum gene manipulation and CC modulation of metabolic activity. Cells that contain the polynucleotides CC of the invention may produce fine chemicals in better yields, with higher CC productivity and/or more efficiently.

XX Sequence 1063 AA;

Query Match 49.5%; Score 266.5; DB 7; Length 1063;
Best Local Similarity 53.6%; Pred. No. 2.3e-22;
Matches 59; Conservative 15; Mismatches 33; Indels 3; Gaps 2;

QY 1 MAADPSTALPDGAGLVLDGVTATLEAGWAKDRIRLOIRKSTGLDVSRRIRVM 60
DB 953 VAANPDSTAGIDGVDGVLVDMEYTERLEAGWAKDRIRLOIRKSTGLDVSRRIRVM 1012

QY 61 SVAEREDWARTRDILAGSILATDFEF-ADLADGVA--IGDGVRSIEK 107
DB 1013 SVPEDEKKEWTTTADHIAEVLATSFETVDALDGETHDIAGVTAKVTK 1062

RESULT 11

ABP66334 standard; protein; 1103 AA.

XX ABP66334;

XX 19-NOV-2002 (first entry)

XX Bifidobacterium longum NCC2705 ORF amino acid sequence SEQ ID NO:1078.

XX Bifidobacterium longum NCC2705; Bifidobacterium; bacterial;

XX antimetabolic; antibacterial; inhibitor of Salmonella; detection;

XX identification; lactic acid bacterium; diarrhoea; pathogenic bacteria;

XX rotavirus; food composition; pharmaceutical composition.

XX Bifidobacterium longum.

XX BP1227152-A1.

XX 31-JUL-2002.

XX 30-JAN-2001; 2001EP-00102050.

XX (NEST) SOC PROD NESTLE SA.

XX WPI; 2002-668397/72.

XX Novel polynucleotide comprising Bifidobacterium genome sequence useful as
XX a probe or primer for detecting and/or identifying Bifidobacterium longum
XX in a biological sample.

XX Claim 3; SEQ ID NO 1078; 80pp; English.

XX The present invention describes a polynucleotide (I) comprising a
XX sequence of a Bifidobacterium genome selected from the nucleotide
XX sequences given in AB081842 and AB081843, or a sequence exhibiting at
XX least 90% identity to which hybridises with the sequences given in
XX AB081842 and AB081843. Also described is a polynucleotide (II) encoding a
XX fusion protein, comprising a sequence selected from 1097 sequences given
XX in ABP6558 to ABP6534 ligated in frame to a polynucleotide encoding a
XX heterologous polypeptide. (I) has antimetabolic and antibacterial
XX activities, and can be used as an inhibitor of Salmonella. (I) (which is

CC a probe) is useful for the detection and/or identification of
CC Bifidobacterium longum in a biological sample. A carrier containing the
CC lactic acid bacterium Bifidobacterium longum NCC2705 (NCIM I-2618) can be
CC used for preventing and/or treating diarrhoea brought about by pathogenic
CC bacteria and/or rotavirus. The carrier is a food composition selected
CC from milk, yogurt, curd, cheese, fermented milk, milk based fermented
CC products, ice-creams, fermented cereal based products, milk based
CC powders, infant formula, pet food or a pharmaceutical composition
CC selected from tablets, liquid bacterial suspensions, dried oral
CC supplement, wet oral supplement, dry tube feeding or wet tube feeding.
CC (I) is useful in DNA arrays or chips to carry out analysis of the
CC expression of the Bifidobacterium gene. AB081844 to AB081850 represent
CC Bifidobacterium related nucleotide sequences given in the Sequence
CC listing from the present invention but not mentioned further within the
CC specification. N.B. The sequence data for this patent is not represented
CC in the printed specification but is based on sequence information
CC supplied by the European Patent Office

XX Sequence 1103 AA;

Query Match 34.4%; Score 185; DB 5; Length 1103;
Best Local Similarity 47.3%; Pred. No. 1.4e-12;
Matches 43; Conservative 17; Mismatches 27; Indels 4; Gaps 2;

QY 8 TAALPDGAGLVLDGVTATLEAGWAKDRIRLOIRKSTGLDVSRRIRVMSVPERE 67
DB 1011 SAALPFG-GFVIIDTADLADLEAGYARDVRSQARKADIDRISLVITVPADV 1069

QY 68 DWARTRDILAGSILATDFEFADLADGVAIG 98

DB 1070 AKYBQFRDLTHTLTATSF--VKGGAELG 1097

RESULT 12

ABU48588 standard; protein; 1091 AA.

XX ABU48588;

XX 19-JUN-2003 (first entry)

XX Protein encoded by Prokaryotic essential gene #34115.

XX Antisense; prokaryotic essential gene; cell proliferation; drug design.

XX Treponema pallidum.

XX WO200277183-A2.

XX 03-OCT-2002.

XX 21-MAR-2002; 2002WO-US009107.

XX 21-MAR-2001; 2001US-00815242.

XX 06-SEP-2001; 2001US-00948993.

XX 25-OCT-2001; 2001US-0342923P.

XX 08-FEB-2002; 2002US-00072851.

XX 06-MAR-2002; 2002US-0362699P.

XX (BLIT-) ELITRA PHARM INC.

XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KU, Zyskind JW;

XX Wall J, Tremick JD, Carr GF, Yamamoto R, Forsyth RA, Xu HH;

XX N-PSDB; ACh52458.

XX WPI; 2003-029926/02.

XX New antisense nucleic acids, useful for identifying proteins or screening
XX PT for homologous nucleic acids required for cellular proliferation to
XX isolate candidate molecules for rational drug discovery programs.
XX Claim 25; SEQ ID NO 76512; 1766pp; English.

CC The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-regulated gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation or to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
CC
CC SQ Sequence 1091 AA;

Query Match 24.5%; Score 132; DB 6; Length 1091;
Best Local Similarity 37.2%; Pred. No. 3.2e-06;
Matches 42; Conservative 16; Mismatches 43; Indels 12; Gaps 5;
DB 6 ESTAALPDGAGLVLDGTVALEABGMAKDRIRLEQLRKSTGLDVSRI--RVMS-- 61
979 ESIAKAINEGTLTVALDTTLTLEDLLEGATIRDLVRGYNLAKERGFSLVDICARVSSDQ 1038
QY -VPAEREDMARTHRDLINAGEIILATDFEPADLADGVAL---GGCV--RVSIIEK 107
DB 1039 DIYCAKRAY-DLHRSYIVGETLAHVQMARVRDASAVVYKSDAVLMEVSIDK 1090

RESULT 13
ADN46385
ID ADN46385 standard; protein; 1065 AA.
XX
AC ADN46385;
XX
DT 01-JUL-2004 (first entry)
XX
DE Thermococcus kodakaraensis KOD1 protein sequence SegID263.
XX
KW gene disruption; gene targeting; marker gene; transformation;
XX homologous recombination; hyperthermostable archaeobacterium; KOD1;
KW gene structure; gene function; enzyme activity; medicine;
KW forensic science; food; drug inspection; molecular biology; immunology.
OS Thermococcus kodakaraensis.
XX
XX MO2004022736-A1.
XX
PD 18-MAR-2004.
XX
PP 29-AUG-2003; 2003MO-IB003597.
XX
PR 30-AUG-2002; 2002JP-00319011.
XX
PA (NISC-) JAPAN SCI & TECHNOLOGY CORP.

XX Imanaka T, Atoml H;
PI WPI; 2004-257583/24.
XX
PT Method for disrupting targeted gene in genome of organism particularly
PT thermostable bacterium and with genome chips for analysis, applicable in
PT studying gene structure and functions.
XX
PS Claim 9; SEQ ID NO 263; 598bp; Japanese.
XX
CC This invention relates to a novel method for targeting disruption of an
CC arbitrary gene in a genome of an organism which comprises providing the
CC whole sequential data of the genome of such organism, selecting at least
CC 1 arbitrary region in the sequence, providing a vector that contains a
CC sequence homologous with the selected region and a marker gene,
CC transformation, and homologous recombination. The genome is preferably
CC the genome of a hyperthermostable archaeobacterium, particularly
CC Thermococcus kodakaraensis KOD1. The method is for targeting the
CC disruption of a gene in the genome of an organism, which is applicable in
CC studying gene structure and functions as well as enzyme activities of
CC encoded proteins and useful in medicine, forensic science, food or drug
CC inspection, molecular biology and immunology. With this method, the
CC disruption of a gene at an arbitrary position in a genome can be achieved
CC efficiently and reliably. The present sequence is that of a protein
CC encoded by the genome of Thermococcus kodakaraensis which was derived
CC using the method of the invention. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
CC
CC SQ Sequence 1065 AA;

Query Match 24.4%; Score 131.5; DB 8; Length 1065;
Best Local Similarity 34.9%; Pred. No. 3.6e-06;
Matches 37; Conservative 16; Mismatches 44; Indels 9; Gaps 2;
QY 11 LPD-----GAGLVLDGTVALEABGMAKDRIRLEQLRKSTGLDVSRI--RVMSVP 63
DB 959 LPDFVABEFDFGVRVVDXTLTLELLAEGLAEFVARIOEMRKRLDLDVNDRIVTIETT 1018
QY 64 AEREDMARTHRDLINAGEIILATDFEPADLADGVALGGCV--RVSIIEK 107
DB 1019 DENRELLQENLDIYMETRAVEFEBKAGYVVEWPEVAKIGIEK 1064

RESULT 14
ABU24074
ID ABU24074 standard; protein; 1035 AA.
XX
AC ABU24074;
XX
DT 19-JUN-2003 (first entry)
XX
DE Protein encoded by prokaryotic essential gene #9601.
XX
KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX
OS Clostridium acetobutylicum.
XX
PN WO200277183-A2.
XX
XX 03-OCT-2002.
XX
PD 21-MAR-2002; 2002MO-US009107.
XX
PP 21-MAR-2001; 2001US-00815242.
XX
PP 06-SEP-2001; 2001US-00948993.
XX
PP 25-OCT-2001; 2001US-0342923P.
XX
PP 08-FEB-2002; 2002US-00072851.
XX
PP 06-MAR-2002; 2002US-0362699P.
XX
PA (ELIT-) ELITRA PHARM INC.

XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zykind JW,
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX WPI: 2003-029926/02.
DR N-PSDB; AKA21944.
XX
PT New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
PS Claim 25; SEQ ID NO 51998; 1766bp; English.
XX
CC The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
XX
SQ Sequence 1035 AA;
Query Match 23.9%; Score 128.5; DB 6; Length 1035;
Best Local Similarity 30.1%; Pred. No. 7.9e-06;
Matches 28; Conservative 26; Mismatches 36; Indels 3; Gaps 1;
QY 18 VLDGTVTATLBAAGAKDRIRELOELKSTGLDVSRIRVMSVAPREEDMARTHDLI 77
DB 941 IVLETTITDREBGEVLRVLSKQNNRKSGFEVNAKIRIYSGNEKLNVRIRKPEDTI 1000
QY 78 AGEILATDFEFADLADGVAL--GDGVRVSIK 107
DB 1001 KETLTATDIYSKNEKKAAYININGEBLNVFVK 1033
RESULT 15
AAU43003
ID AAU43003 standard; protein; 1185 AA.
XX
XX AAU43003;
XX
XX 27-FEB-2002 (first entry)
DB Propionibacterium acnes immunogenic protein #3899.
XX
XX
XX SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;

KW dermatological; osteopathic; neuroprotectant.
XX
XX
OS Propionibacterium acnes.
XX
XX W0200181581-A2.
XX
XX
XX 01-NOV-2001.
XX
XX 20-APR-2001; 2001WO-US012865.
XX
XX 21-APR-2000; 2000US-0199047P.
XX 02-JUN-2000; 2000US-0208841P.
XX 07-JUL-2000; 2000US-0216747P.
XX
XX (CORI-) CORIXA CORP.
XX
XX
PI Skeiky YW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
PI L'valsomneuve J, Zhang Y, Jen S, Carter D;
XX
XX WPI: 2001-616774/71.
DR N-PSDB; AAS59519.
XX
XX
PT Propionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for
PT treating acne vulgaris.
XX
XX
PS Example 1; SEQ ID NO 4198; 1069pp; English.
XX
CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
CC polypeptides. The proteins and their associated DNA sequences are used in
CC the treatment, prevention and diagnosis of medical conditions caused by
CC *P. acnes*. The disorders include SAPHO syndrome (synovitis, acne,
CC pustulosis, hyperostosis and osteomyelitis), uveitis and endophthalmitis.
CC *P. acnes* is also involved in infections of bone, joints and the central
CC nervous system, however it is particularly involved in the inflammatory
CC lesions associated with acne vulgaris. A method for detecting the
CC presence or absence of *P. acnes* in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies
CC specific for *P. acnes* proteins. These antibodies can be used to
CC downregulate expression and activity of *P. acnes* polypeptides and
CC therefore treat *P. acnes* infections. The antibodies may also be used as
CC diagnostic agents for determining *P. acnes* presence, for example, by
CC enzyme linked immunosorbent assay (ELISA). Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
XX
SQ Sequence 1185 AA;
Query Match 23.7%; Score 127.5; DB 4; Length 1185;
Best Local Similarity 27.9%; Pred. No. 1.3e-05;
Matches 41; Conservative 18; Mismatches 47; Indels 41; Gaps 4;
QY 2 AADPESTAA-----LPD---GAGLVVDGTVTATLBAAGW----- 33
DB 1038 AADPEMLASLAKSGVEMDVFEVGEKAVVTADVIYSRPREGMSVYNGEFTVALDL 1097
QY 34 -----AKDRIRELOELKSTGLDVSRIRVMSVAPREEDMARTHDLIAGEIL 82
DB 1098 EITPELBAAGQAREVIRFVQDSRKKAGLDVSDRITTLAWSASDLATVAIEBHAIQISEVL 1157
QY 83 ATDFEFADLADGVALGD--GVRVSIK 107
DB 1158 AVQMSRBRPADDWAVEPDGLAVKVK 1184
/ Search completed: April 14, 2006, 17:18:47
Job time : 43.8457 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: April 14, 2006, 17:19:08 ; Search time 6.33288 Seconds
(without alignments)
1640.866 Million cell updates/sec

Title: US-10-620-246-4

Perfect score: 538

Sequence: 1 MAADPESTALPDAGLVVL.....ADLADGVALIGDGVRSIEKT 108

Scoring table: BLOSUM62
Gapop 10.0 , Gapept 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	534	99.3	1041	2	E70760	probable 11es prot
2	423	78.6	1059	2	E87058	isooleucyl-tRNA syn
3	174.5	32.4	1078	2	E75407	isooleucyl-tRNA syn
4	139.5	25.9	1047	2	T34946	probable isooleucyl
5	132	24.5	1091	2	E71322	isooleucine-tRNA 11
6	128.5	23.9	1035	2	G97273	isooleucyl-tRNA syn
7	126.5	23.5	1066	2	F71100	isooleucine-tRNA 11
8	123.5	23.0	1067	2	H75139	isooleucyl-tRNA syn
9	111	20.6	1044	2	H69049	isooleucine-tRNA 11
10	110	20.4	1045	1	SYX1	isooleucine-tRNA 11
11	105	19.5	1036	2	F71565	probable isooleucin
12	104.5	19.4	986	2	E90220	isooleucine-tRNA sy
13	101.5	18.9	1064	2	T40751	isooleucyl-tRNA syn
14	100.5	18.7	1070	2	H84369	isooleucyl-tRNA syn
15	97.5	18.1	1036	2	F81719	isooleucine-tRNA 11
16	96	17.8	1072	1	SYB14	isooleucine-tRNA 11
17	95.5	17.8	1064	2	E72729	probable isooleucyl
18	94.5	17.6	1106	2	A97819	isooleucine-tRNA 11
19	93	17.3	1042	2	H70203	isooleucine-tRNA 11
20	89	16.5	1086	2	B71667	isooleucine-tRNA 11
21	88	16.1	1018	2	A69339	probable isooleucin
22	86.5	16.1	1039	2	C64418	isooleucine-tRNA 11
23	82.5	15.3	659	2	D84286	3-hydroxyacyl-CoA
24	82	15.2	332	2	AH1594	hypochemical prote
25	80.5	15.0	309	2	G98311	hypochemical prote
26	80.5	15.0	309	2	AC2971	hypochemical prote
27	79	14.7	375	2	G70966	isooleucine-tRNA 11
28	79	14.7	485	2	T01968	isooleucine-tRNA 11
29	79	14.7	1254	2	T04047	isooleucine-tRNA 11

30	77.5	14.4	569	2	T35476	probable regulator
31	77	14.3	8563	2	T30226	polyketide synthas
32	76.5	14.2	431	2	A59474	tRNA adenylyltrans
33	76.5	14.2	662	2	I38400	melanoma-associate
34	76.5	14.2	668	2	A41234	melanocyte-specifi
35	76.5	14.2	1199	1	S76549	transcription-repa
36	76	14.1	432	2	T35263	probable carboxype
37	76	14.1	1266	2	I59314	isooleucine-tRNA 11
38	75.5	14.0	449	2	B75451	NADH oxidase - Del
39	75	13.9	441	2	T35788	probable transcrip
40	74	13.8	368	2	T35150	probable glycosyl
41	73.5	13.7	393	2	S38875	methionine adenosy
42	73.5	13.7	396	2	A70581	hypochemical prote
43	72.5	13.5	256	2	H83002	indazoleglycerol-
44	72.5	13.5	301	2	A32498	Mx resistance prot
45	72.5	13.5	435	2	C69194	L-asparaginase I -

ALIGNMENTS

RESULT 1

E70760 Probable 11es protein - Mycobacterium tuberculosis (strain H37Rv)

C/Species: Mycobacterium tuberculosis

C/Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004

C/Accession: E70760

R/Colo, S.T.; Broach, R.; Parkhill, J.; Garner, T.; Churcher, C.; Harris, D.; Gordon, J.; Connor, R.; Davies, R.; Devlin, K.; Felwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.

Nature 393, 537-544, 1998

A/Authors: Sgares, R.; Sulston, J.B.; Taylor, K.; Whitehead, S.; Barrett, B.G.

A/Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome

A/Reference number: A70500; PMID:9825987; PMID:9634230

A/Accession: E70760

A/Status: preliminary; nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA

A/Residues: 1-1041 <COL>

A/Cross-references: UNIPROT:Q10765; UNIPARC:UP100001364FA; GB:274020; GB:AL123456; NID::

A/Experimental source: strain H37Rv

A/Genetic:

A/Gene: 11es

C/Superfamily: isooleucine-tRNA ligase

Query Match 99.3%; Score 534; DB 2; Length 1041;

Best Local Similarity 99.1%; Pred. No. 9.2e-43; Matches 107; Mismatches 0; Indels 0; Gaps 0;

Matches 107; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAADPESTALPDAGLVVDGVTVALEAGWAKRIRRELQIRKSTGLDVSRRIVM 60

Db 934 VAADPESTALPDAGLVVDGVTVALEAGWAKRIRRELQIRKSTGLDVSRRIVM 993

QY 61 SVPAERDWARTRDLIAGIILATDFEFADLADGVALIGDGVRSIEKT 108

Db 994 SVPAERDWARTRDLIAGIILATDFEFADLADGVALIGDGVRSIEKT 1041

QY 87058

isooleucyl-tRNA synthase [Imported] - Mycobacterium leprae

C/Species: Mycobacterium leprae

C/Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004

C/Accession: E87058

R/Colo, S.T.; Bismeyer, K.; Parkhill, J.; James, K.D.; Thomson, N.R.; Wheeler, P.R.; H.

eam, M.A.; Rutherford, K.M.

Nature 409, 1007-1011, 2001

A/Authors: Rutter, S.; Seeger, K.; Simon, S.; Simmonds, M.; Skelton, J.; Squares, R.; S.

A/Title: Massive gene decay in the leprosy bacillus.

A/Reference number: A86909; PMID:21128732; PMID:11234002

A/Accession: E87058

A/Status: preliminary

A/Molecule type: DNA

	Matches	28;	Conservative	26;	Mismatches	36;	Indels	3;	Gaps	1;			
Qy	18	VULDGVTATLEAEGMAKDRIRRELOELAKSTGLDVSRRVMSVAEREDMARTRDLI	77	:	: :	:	:	:	:	:			
Dd	941	IYLETITTDIREGGYLAEVLISKVNMRKSSEFEVDADKLIIYSNGEIKLENVIKKPDDTI	1000	:	:	:	:	:	:	:			
Qy	78	AGEIILATDFEPADLADVAI---	GDGVRSIER	107	:	:	:	:	:	:			
Dd	1001	KKETLATDIISSENKEALIVNGBELNVPAVK	1033	:	:	:	:	:	:	:			
RESULT 7 F71100													
Isoleucine-tRNA ligase (EC 6.1.1.5) - Pyrococcus horikoshii													
CSpecies:	Pyrococcus horikoshii												
CDate:	14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 05-Oct-2004												
AAccession:	F71100												
R.Kawarabayashi,	Y.; Sawada,	M.; Horikawa,	H.; Hino,	Y.; Yamamoto,	S.; Sekina,	M.; Ohfuku,	Y.; Funahashi,	T.; Tanaka,	T.; Kudoh,	Y.; Yamazaki,	J.; Kushida,	N.; Oguchihara,	E.
DNA Res.	5, 55-76, 1998												
AArticle:	Complete sequence and gene organization of the genome of a hyper-thermophilic archaeon from Japan												
AReference number:	A71000; PMID:98344137; PMID:9679194												
AStatus:	preliminary; nucleic acid sequence not shown; translation not shown												
A.Molecule type:	DNA												
A.Residues:	1-1066 <KAN>												
A.Cross-references:	UNIPROT:O58792; UNIPARC:UPI00001364FE; GB:APO00004; NID:g3236131; PDB:1WUJ												
A.Experimental source:	strain OT3												
A.Note:	this accession replaces an interim accession for a sequence replaced by GenBank C:Genbank:PF1065												
A.Gene:	PHI065												
C.Superfamily:	isoleucine-tRNA ligase												
C.Keywords:	aminoacyl-tRNA synthetase; ligase; protein biosynthesis												
<hr/>													
Query Match	Best Local Similarity	23.5%; Score 126.5; DB 2; Length 1066; Matches 36; Conservative 17; Mismatches 44; Indels 9; Gaps 2;											
Qy	11	LDP-----GAGLVLDGTVALLEAEGMAKDRIRRELOELAKSGLDVSDRIRVMASVP	63	:	:	:	:	:	:	:			
Dd	959	LEDPLAVADEPEGGRVVYDKLTTLRELIAELGLEFRRIQENMKRLDLVDNDRIVTTIEPT	1018	:	:	:	:	:	:	:			
Qy	64	AAREDMARKTRDLIAGETLATDFEFADLADGAIGDY--RVSIKR	107	:	:	:	:	:	:	:			
Dd	1019	DNRRILLQENTLDYMRETTRAILEVRFBKAQGYEWPEYQAOKIGIKR	1064	:	:	:	:	:	:	:			
RESULT 8 H75139													
Isocytosyl-tRNA synthetase (Iles) PAB0616 - Pyrococcus abyssi (strain Orsay)													
CSpecies:	Pyrococcus abyssi												
CDate:	20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004												
AAccession:	H75139												
RAnonymous,	Genoscope												
A.Submitted to the EMBL Data Library,	July 1999												
A.Description:	Pyrococcus abyssi genome sequence: insights into archeal chromosome stru-												
A.Reference number:	A75001												
A.Accession:	H75139												
A.Status:	preliminary												
A.Molecule type:	DNA												
A.Residues:	1-1067 <KAN>												
A.Cross-references:	UNIPROT:Q9V073; UNIPARC:UPI00000344D8; GB:NJ248285; GB:AL096836; NTID:M75139												
A.Experimental source:	Strain Orsay												
C.Genetics:													
A.Gene:	Iles; PAB0616												
C.Superfamily:	isoleucine-tRNA ligase												
<hr/>													
Query Match	Best Local Similarity	23.0%; Score 123.5; DB 2; Length 1067; Matches 39; Conservative 14; Mismatches 39; Indels 19; Gaps 3;											
Qy	11	LDP-----GAGLVLDGTVALLEAEGMAKDRIRRELOELAKSGLDVSDRIRVMASVP	63	:	:	:	:	:	:	:			

Db 959 PPDFVSBEPFGGVFDKTLTRELIALGALREFFVRIOEWRKRLDLDVNDRIIVTIEET 1018

QY 64 AREDMARTHDLLAGELIATDFEPADLADGVAIGDGY-----RSTIEK 107

Db 1019 DENVELLKENDLYIMRETRADKIVF-----GKAGYVWEPEVQAKIGIER 1064

RESULT 9

HE9049
Isoleucine-tRNA ligase (EC 6.1.1.5) - Methanobacterium thermoautotrophicum (strain Delta H)
N:Alternate names: Isoleucyl-tRNA synthetase
C:Species: Methanobacterium thermoautotrophicum
C:date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
C:Accession: H69049
R:Smith, D.R.; Doucette-Stamm, L.A.; Delouhery, C.; Lee, H.; Dubois, J.; Aldredge, T.; Qiu, D.; Spadafora, R.; Vicaltre, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jivan, N. K.; S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
J: Bacteriol. 179, 7135-7155, 1997
A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: func
A:Reference number: A69000; MUID:98037514; PMID:9371463
A:Accession: H69049
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-1044 <MTH>
A:Cross-references: UNIPROT:O27428; UNIPARC:UP10000165BB4; GB:AB000900; GB:AB000666; N
C:GeneticB:
A:Experimental source: strain Delta H

C:Keywords: aminoacyl-tRNA synthetase; ligase; protein biosynthesis

QY Query Match 20.6%; Score 111; DB 2; Length 1044;
Best Local Similarity 28.1%; Pred. No. 0.017;
Matches 25; Conservative 21; Mismatches 41; Indels 2; Gaps 1;

Db 934 PENTVAAPFDGGSVFDTLTPETMSAMRELVRIQDKOLDLDEASIEVSVCSE 993

QY 65 EREDMARTHDLLAGELIAT--DFEPADL 91

Db 994 ERELTEPQREFIENEVRASITLSPDSESL 1022

RESULT 10

SYN1
Isoleucine-tRNA ligase (EC 6.1.1.5) [validated] - Methanobacterium thermoautotrophicum
N:Alternate names: Isoleucyl-tRNA synthetase
C:Species: Methanobacterium thermoautotrophicum
C:date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 03-Jun-2002
C:Accession: A40398
R:Jenal, U.; Recheleiner, T.; Tan, P. Y.; Buehlmann, E.; Meile, L.; Leisinger, T.
J: Biol. Chem. 266, 10570-10577, 1991
A:Title: Isoleucyl-tRNA synthetase of Methanobacterium thermoautotrophicum Marburg. Clon
id.
A:Reference number: A40398; MUID:91244836; PMID:2037598
A:Accession: A40398
A:Molecule type: DNA
A:Residues: 1-1045 <JEN>
A:Cross-references: UNIPARC:UP1000016FBA0; GB:H59245; NID:g149726; PIDN:AAA72950.1; PID
C:Function:
A:Description: EC 6.1.1.5 [validated, MUID:91244836]
C:Superfamily: Isoleucine-tRNA ligase
C:Keywords: aminoacyl-tRNA synthetase; ATP; ligase; protein biosynthesis

QY Query Match 20.4%; Score 110; DB 1; Length 1045;
Best Local Similarity 28.1%; Pred. No. 0.021;
Matches 25; Conservative 20; Mismatches 42; Indels 2; Gaps 1;

Db 934 PENTVAAPFDGGSVFDTLTPETMSAMRELVRIQDKOLDLDEARIEVSVCSP 993

QY 65 EREDWARTHRDLIAGEIAT--DPEPADL 91
 Db 994 ERELETPQREFEVNEVRASHLSFDYTEL 1022

RESULT 11

F71565
 probable isoleucine-tRNA ligase (EC 6.1.1.5) - Chlamydia trachomatis (serotype D, strain C/Species: Chlamydia trachomatis
 C/Date: 13-Sep-1998 #sequence_revision 13-Sep-1998 #text_change 09-Jul-2004
 C/Accession: F71565
 R/Stephen, R.S.; Kalman, S.; Lammel, C.J.; Fan, J.; Marche, R.; Aravind, L.; Mitchell, Science 282, 754-759, 1998
 A/Title: Genome sequence of an obligate intracellular pathogen of humans: Chlamydia trachomatis
 A/Reference number: A71570; M01D:99000809; PMID:9784136
 A/Accession: F71565
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-1036 <ARN>
 A/Cross-references: UNIPROT:O84022; UNIPARC:UPI00001364EB; GB:AE001277; GB:AE001273; NID: A/Experimental source: serotype D, strain UW-3/Cx
 C/Genetic:
 A/Genes: 11es
 C/Superfamily: isoleucine-tRNA ligase
 C/Keywords: ligase

Query Match 19.5%; Score 105; DB 2; Length 1036;
 Best Local Similarity 28.3%; Pred. No. 0.062;

Matches 26; Conservative 18; Mismatches 48; Indels 0; Gaps 0;

QY 3 ADESTALPDGALVLDGTVTAELAEAGMADRIRELOELKSTGLDVSRIYVMV 62

Db 926 ASAEGRVARSASFVAVLDQCLTEPLIMEGIARELVKIMWRNRRLAHVSDRIARLHA 985

QY 63 PAEREDWARTHRDLIAGEIATDPEPADLADG 94

Db 986 PVIYQEAFLAKKEYICEETLTTSVSYVDYKEG 1017

RESULT 12

E90220
 Isoleucine-tRNA synthetase (11es) [imported] - Sulfolobus solfataricus
 C/Species: Sulfolobus solfataricus
 C/Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 09-Jul-2004
 C/Accession: E90220
 R/She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan- arrett, R.A.; Ragan, M.A.; Sersen, C.W.; Van der Oost, J.
 submitted to Genbank, April 2001
 A/Description: Sulfolobus solfataricus complete genome.
 A/Reference number: A99139
 A/Accession: E90220
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-986 <RUR>
 A/Cross-references: UNIPROT:Q9UXB1; UNIPARC:UPI0000066A25; GB:AE006641; NID:G13813890; F C/Genetic:
 A/Genes: 11es
 C/Superfamily: isoleucine-tRNA ligase

Query Match 19.4%; Score 104.5; DB 2; Length 986;
 Best Local Similarity 27.6%; Pred. No. 0.065;

Matches 27; Conservative 20; Mismatches 46; Indels 5; Gaps 1;

QY 16 GLVYLDGTVTAELAEAGMADRIRELOELKSTGLDVSRIYVMVPAEREDWARTHRD 75

Db 888 GIYVISEKEISEESEGRLRIIRIQFMKQKLANLDYIEISMKVPEERVTITQWES 947

QY 76 LIAGEIATDPEPADLADGVAI-----GDGVRVSIERT 108

Db 948 FIKSETRASNIIGBAKGITMDMDIGESYITIGIKS 985

RESULT 13

T40751
 isoleucyl-tRNA synthetase, cytoplasmic - fission yeast (Schizosaccharomyces pombe)
 C/Species: Schizosaccharomyces pombe
 C/Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
 C/Accession: T40751
 R/Lyne, M.; Rajadream, M.A.; Barrell, B.G.; Lauber, J.; Hilbert, H.; Duesterhoeft, A.
 submitted to the EMBL Data Library, March 1998
 A/Reference number: Z21948
 A/Accession: T40751
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-1064 <LYN>
 A/Cross-references: UNIPROT:O13651; UNIPARC:UPI00001364B3; EMBL:AL022072; PIDN:CAA17821
 A/Experimental source: strain 972h; cosmid c8D2
 C/Genetic:
 A/Genes: p1058; SPDB:SPBC8D2.06
 A/Map position: 2
 C/Superfamily: isoleucine-tRNA ligase

Query Match 18.9%; Score 101.5; DB 2; Length 1064;
 Best Local Similarity 30.4%; Pred. No. 0.14;

Matches 28; Conservative 16; Mismatches 23; Indels 25; Gaps 2;

QY 13 DGAALVLDGTVTAELAEAGMADRIRELOELKSTGLDVSRIYV----- 58

Db 958 DGIYVLDIIEIDAOGLAEGLAREVINRVQRKRKSNLQVTDVMTYKIKNDITGLBSA 1017

QY 59 -----VMSVPAEREDWARTHRDLADG 80

Db 1018 VDSNEALFSKVLRRPIEKETGA-DESNITASE 1048

RESULT 14

H84369
 isoleucyl-tRNA synthetase [imported] - Halobacterium sp. NRC-1
 C/Species: Halobacterium sp. NRC-1
 C/Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
 C/Accession: H84369
 R/Ng, W.V.; Kennedy, S.P.; Mahatras, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Laeky, J.; Leithauer, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jabl, Jung, K.H.; Alam, M.; Freltas, T.
 Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
 A/authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; L A/Title: Genome sequence of Halobacterium species NRC-1.
 A/Reference number: A84160; M01D:20504483; PMID:11016950
 A/Accession: H84369
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-1070 <STC>
 A/Cross-references: UNIPROT:Q9HN97; UNIPARC:UPI0000063ABD; GB:AE004437; NID:G10581609; F C/Genetic:
 A/Genes: 11es
 C/Superfamily: isoleucine-tRNA ligase

Query Match 18.7%; Score 100.5; DB 2; Length 1070;
 Best Local Similarity 33.3%; Pred. No. 0.17;

Matches 27; Conservative 17; Mismatches 34; Indels 3; Gaps 2;

QY 15 AGIYVLDGTVTAELAEAGMADRIRELOELKSTGLDVSRIYVMVSPAER-EDWARTH 73

Db 973 AGVYVDTIELNEVSEEGYAREVRRQERKEMDLAMDAEIRLDLVFDERVGLVARH 1032

QY 74 RDLIAGEIATDPEPADLADG 94

Db 1033 EPLITAEPRAR--ELGEVEDG 1051

RESULT 15

F81719
 isoleucyl-tRNA synthetase TC0288 [imported] - Chlamydia muridarum (strain N199)
 C/Species: Chlamydia muridarum, Chlamydia trachomatis MoPn
 C/Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004

C/Accession: F81719
R/Read, T.D.; Brunham, R.C.; Shen, C.; Gail, S.R.; Heidelberg, J.F.; White, O.; Hickey,
C.; Dodson, R.; Gwin, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg,
Nucleic Acids Res. 28, 1397-1406, 2000
A/Title: Genome sequences of Chlamydia trachomatis Mopn and Chlamydia pneumoniae AR39.
A/Reference number: A81500; MUID:20150255; PMID:10684935
A/Accession: F81719
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-1036 <TET>
A/Cross-references: UNIPROT:Q9PL20; UNIPARC:UPI0000057880; GB:AB002296; GB:AB002160; NIT
A/Experimental source: strain N199 (Mopn)
C/Genetics:
A/Gene: TC0288
C/Superfamily: Isoleucine-tRNA ligase

Query Match 18.1%; Score 97.5; DB 2; Length 1036;
Best Local Similarity 28.0%; Pred. No. 0.32; Indels 3; Gaps 1;
Matches 26; Conservative 21; Mismatches 43;

QY 17 LVVLDTVTALAEAGWAKDIRRELQRIKSTGLDVSDRIRVWVSVPAREDMARTHDL 76
DB 940 VALIDCQLTSPLEMEGIAREIVKINTMRKNGKLVSDRIAIRLHAPKIVQEARFQYBEY 999

QY 77 IAGEIILATDPEFADLADGV--AIGDGVRSIE 106
DB 1000 ICEETLTTSVSFTIDKEGBEWDVNGHAVSLSE 1032

Search completed: April 14, 2006, 17:34:35
Job time : 7.33288 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: April 14, 2006, 17:04:10 ; Search time 39.8484 Seconds
(without alignments)
1912.171 Million cell updates/sec

Title: US-10-620-246-4

Perfect score: 538

Sequence: 1 MADDPSTALPDAGLVL.....ADLADGVALGDGVRVSIKTK 108

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : UniProt_05.80.*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	534	99.3	1041	1 SYI_MYCTU	Q10765 mycobacteri
2	534	99.3	1041	2 Q7VEZ0_MYCBO	Q7VEZ0 mycobacteri
3	454	84.6	1053	2 Q740U6_MYCPA	Q740U6 mycobacteri
4	423	78.6	1059	1 SYI_MYCLE	Q9X755 mycobacteri
5	352.5	65.5	1042	2 Q5YTW9_NOCRA	Q5YTW9 nocardia fa
6	277.5	51.6	1103	2 Q8FNV0_CORER	Q8FNV0 corynebacte
7	270.5	50.3	1092	2 Q4JWB5_CORJK	Q4JWB5 corynebacte
8	266.5	49.5	1052	2 Q6NNDP_CORDI	Q6NNDP corynebacte
9	266.5	49.5	1054	2 Q8NNP0_CORGL	Q8NNP0 corynebacte
10	197.5	36.7	1091	2 Q4H7J7_GDEIC	Q4H7J7 bifidobacte
11	185	34.4	1103	2 Q8G3J2_BIFLO	Q8G3J2 bifidobacte
12	174.5	32.4	1078	2 Q9RUP8_DEIRA	Q9RUP8 deinococcus
13	167	31.0	1058	2 Q8TNE2_METRA	Q8TNE2 methanosaar
14	158.5	29.5	1058	2 Q9P9L9_METRA	Q9P9L9 methanosaar
15	156	28.0	1058	2 Q8PSV9_METWA	Q8PSV9 methanosaar
16	152.5	28.3	1074	2 Q5UZY4_HALMA	Q5UZY4 halorocula
17	148.5	27.6	1109	2 Q4NBB3_SMICC	Q4NBB3 archaobacte
18	147.5	27.4	1047	2 Q8ZAC9_STRAM	Q8ZAC9 streptomyc
19	139.5	25.9	1047	2 Q9S2A5_STRCO	Q9S2A5 streptomyc
20	139.5	25.9	1162	2 Q8A9K9_BACTN	Q8A9K9 bacteroides
21	138.5	25.7	1213	2 Q7UNZ2_RHOBA	Q7UNZ2 rhodospirill
22	137.5	25.6	1137	2 Q7MUD3_PORGI	Q7MUD3 porphyromon
23	134.5	25.0	1141	2 Q5LCU8_BACFR	Q5LCU8 bacteroides
24	134.5	25.0	1141	2 Q64U07_BACFR	Q64U07 bacteroides
25	132	24.5	1091	1 SYI_TREBP	Q83466 treponema p
26	131.5	24.4	1065	2 Q5J3J1_PYRKO	Q5J3J1 pyrococcus
27	130	23.9	1033	2 Q63BZ8_BACCF	Q63BZ8 bacillus ce
28	128.5	23.9	1035	2 Q97BS0_CLOAB	Q97BS0 clostridium
29	128.5	23.9	1043	1 SYI_THET8	Q56650 thermus the
30	128.5	23.9	1067	2 Q72JN6_THET2	Q72JN6 thermus the
31	128.5	23.9	1084	2 Q8KFL5_CHLME	Q8KFL5 chlorobium

32	128	23.8	1033	2 Q739A1_BACCI	Q739A1 bacillus ce
33	127.5	23.7	1114	2 Q6AB89_PROAC	Q6AB89 propionibac
34	126.5	23.5	1066	1 SYI_PYRHO	Q58792 pyrococcus
35	126	23.4	1033	2 Q4MHJ2_BACCS	Q4MHJ2 bacillus ce
36	126	23.4	1033	2 Q6HUF2_BACHK	Q6HUF2 bacillus th
37	126	23.4	1033	2 Q81R75_BACAN	Q81R75 bacillus an
38	126	23.4	1056	2 Q831F6_TROW8	Q831F6 tropheryma
39	126	23.4	1066	1 SYI_PYRPU	P46214 pyrococcus
40	126	23.4	1066	2 Q83G52_TROWT	Q83G52 tropheryma
41	125.5	23.3	1035	2 Q81E30_BACCR	Q81E30 bacillus ce
42	123.5	23.0	1067	1 SYI_PYRAB	Q9V072 pyrococcus
43	119	22.1	1111	2 Q73HW7_WOLPM	Q73HW7 wolbachia p
44	118.5	22.0	1039	2 Q8XHS4_CLOES	Q8XHS4 clostridium
45	118.5	22.0	1119	2 Q6AF20_LEIXX	Q6AF20 leifsonia x

ALIGNMENTS

RESULT 1
SYI_MYCTU STANDARD; PRT; 1041 AA.
ID SYI_MYCTU
AC Q10765; 006181;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Isolation of Mycobacterium tuberculosis complex (EC 6.1.1.5) (Isolation of Mycobacterium tuberculosis complex)
DE (118S).
GN Name: 118S, Ordered locus names: RV1536, MT1587; ORF names: MYC148.29C;
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteriales; Actinomycetales;
OC Corynebacteriaceae; Mycobacteriaceae; Mycobacterium;
OC Mycobacterium tuberculosis complex.
OC NCBI_TaxID=1773;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=H37RV;
RX MEDLINE=98295987; PubMed=9634230; DOI=10.1038/31159;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C.M.,
RA Harris D.B., Gordon S.V., Eigmeier K., Gas S., Barry C.B. III,
RA Tekala F., Badcock K., Basham D., Brown D., Chillingworth T.,
RA Connor R., Davies R.M., Devlin K., Fellwell T., Gentles S., Hamlin N.,
RA Holtroft S., Hornby T., Jagels K., Krogh A., McLean J., Moulie S.,
RA Murphy L.D., Oliver S., Osborne J., Quail M.A., Rajadream M.A.,
RA Rogers J., Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sultson J.B., Taylor K., Whitehead S., Barrall B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence."
RL Nature 393:537-544(1998).
RP [2]
RN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=CDC 1551 / OshKob.
RX MEDLINE=22206494; PubMed=12218036;
RA DOI=10.1128/JB.184.19.5479-5490.2002;
RA Pletschmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J.D., Deboy R.T., Dodson R.J., Gaitan M.L., Haft D.H.,
RA Hickey E.K., Kolonay J.F., Nelson W.C., Unanue L.A., Ermolenko M.D.,
RA Salzberg S.L., Delcher A., Uterback T.R., Weidman J.F., Kouri H.M.,
RA Gill J., Mikula A., Bishal W., Jacobs W.R. Jr., Venter J.C.,
RA Fraser C.M.;
RT "Whole-genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains."
RL J. Bacteriol. 184:5479-5490(2002).
CC -I- CANNARYTIC ACTIVITY: ATP + L-isoleucine + tRNA(Ile) = AMP +
CC dihydrophosphate + D-isoleucyl-tRNA(Ile).
CC -I- COFACTOR: Binds 1 zinc ion per subunit (by similarity).
CC -I- SUBCELLULAR LOCATION: Cytoplasmic.
CC -I- SIMILARITY: Belongs to the class-I aminoacyl-tRNA synthetase
CC family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

CC use as long as its content is in no way modified and this statement is not removed.

CC EMBL: BX842577; CA98326.1; -; Genomic DNA.
 CC EMBL: AE000516; AK45854.1; -; Genomic DNA.
 CC PIR: E70760; E70760.
 DR HSSP; P56690; 11LE.
 DR TIGR; MT1587; -.
 DR Tuberculat; RV1536; -.
 DR InterPro; IPR002300; tRNA-synt_1a.
 DR InterPro; IPR001412; tRNA-synt_1.
 DR InterPro; IPR002301; tRNA-synt_1le.
 DR Pfam; PF00133; tRNA-synt_1; 1.
 DR PRINTS; PR00984; TRNASYNTHLE.
 DR TIGRFAm; TIGR00392; 11es; 1.
 DR PROSITE; PS00178; AA_trna_ligase_1; 1.
 KW Aminocyl-tRNA synthetase; ATP-binding; Complete proteome; Ligase;
 KW Metal-binding; Nucleotide-binding; Protein biosynthesis; Zinc.
 FT MOTIF 53 63 "KMSKS" region.
 FT BINDING 619 623 ATP (By similarity).
 FT SEQUENCE 1041 AA; 117340 MW; B5023822848E08C6 CRC64;

Query Match 99.3%; Score 534; DB 1; Length 1041;
 Best Local Similarity 99.1%; Pred. No. 1.3e-39;
 Matches 107; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAADPESTALPDGAGLVLDGVTATLEAEAGAKXRIEQLRKSTGLDVSDRIRVVM 60
 ID Q740U6 MYCBO PRELIMINARY; PRT; 1041 AA.
 AC Q740U6;
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DR Isoeucyl-tRNA synthetase 11es (EC 6.1.1.5).
 GN Name=11es; OrderedlocusNames=Mb1563;
 OS Mycobacterium bovis;
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacteriaceae; Mycobacterium;
 OC Mycobacterium tuberculosis complex.
 OX NCBI_TaxID=1765;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=AF2122/97;
 RX MEDLINE=22709107; PubMed=12788972; DOI=10.1073/pnas.1130426100;
 RA Garner T., Elsigmeier K., Camus J.-C., Medina N., Mansoor H.,
 RA Pryor M., Duthey S., Grondin S., Lacroix C., Monsemp C., Simon S.,
 RA Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,
 RA Barthill J., Barrell B.G., Cole S.T., Gordon S.V., Hewinson R.G.,
 RT "The complete genome sequence of Mycobacterium bovis";
 RL Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882 (2003).
 DR EMBL; BX248339; CAD96230.1; -; Genomic DNA.
 DR HSSP; P56690; 11LE.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0004822; F:isoleucine-tRNA ligase activity; IEA.
 DR GO; GO:0016874; F:Ligase activity; IEA.
 DR GO; GO:0006428; P:isoleucyl-tRNA aminoacylation; IEA.
 DR InterPro; IPR002300; tRNA-synt_1a.
 DR InterPro; IPR001412; tRNA-synt_1.
 DR InterPro; IPR002301; tRNA-synt_1le.
 DR Pfam; PF00133; tRNA-synt_1; 1.
 DR PRINTS; PR00984; TRNASYNTHLE.
 DR TIGRFAm; TIGR00392; 11es; 1.
 DR PROSITE; PS00178; AA_trna_ligase_1; UNKNOWN_1.

RESULT 2
 Q740U6 MYCBO PRELIMINARY; PRT; 1041 AA.

ID Q740U6 MYCBO PRELIMINARY; PRT; 1041 AA.
 AC Q740U6;
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DR Isoeucyl-tRNA synthetase 11es (EC 6.1.1.5).
 GN Name=11es; OrderedlocusNames=Mb1563;
 OS Mycobacterium bovis;
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacteriaceae; Mycobacterium;
 OC Mycobacterium tuberculosis complex.
 OX NCBI_TaxID=1765;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=AF2122/97;
 RX MEDLINE=22709107; PubMed=12788972; DOI=10.1073/pnas.1130426100;
 RA Garner T., Elsigmeier K., Camus J.-C., Medina N., Mansoor H.,
 RA Pryor M., Duthey S., Grondin S., Lacroix C., Monsemp C., Simon S.,
 RA Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,
 RA Barthill J., Barrell B.G., Cole S.T., Gordon S.V., Hewinson R.G.,
 RT "The complete genome sequence of Mycobacterium bovis";
 RL Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882 (2003).
 DR EMBL; BX248339; CAD96230.1; -; Genomic DNA.
 DR HSSP; P56690; 11LE.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0004822; F:isoleucine-tRNA ligase activity; IEA.
 DR GO; GO:0016874; F:Ligase activity; IEA.
 DR GO; GO:0006428; P:isoleucyl-tRNA aminoacylation; IEA.
 DR InterPro; IPR002300; tRNA-synt_1a.
 DR InterPro; IPR001412; tRNA-synt_1.
 DR InterPro; IPR002301; tRNA-synt_1le.
 DR Pfam; PF00133; tRNA-synt_1; 1.
 DR PRINTS; PR00984; TRNASYNTHLE.
 DR TIGRFAm; TIGR00392; 11es; 1.
 DR PROSITE; PS00178; AA_trna_ligase_1; UNKNOWN_1.

KW Aminocyl-tRNA synthetase; Complete proteome; Ligase.
 KW SEQUENCE 1041 AA; 117313 MW; 451338283248D3AB CRC64;

Query Match 99.3%; Score 534; DB 2; Length 1041;
 Best Local Similarity 99.1%; Pred. No. 1.3e-39;
 Matches 107; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAADPESTALPDGAGLVLDGVTATLEAEAGAKXRIEQLRKSTGLDVSDRIRVVM 60
 ID Q740U6 MYCBA PRELIMINARY; PRT; 1053 AA.
 AC Q740U6;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DR Isoeucyl-tRNA synthetase 11es (EC 6.1.1.5).
 GN Name=11es; OrderedlocusNames=MAP1246;
 OS Mycobacterium paratuberculosis.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacteriaceae; Mycobacterium;
 OC Mycobacterium avium complex (MAC).
 OX NCBI_TaxID=1770;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=R10;
 RA Li L., Bannantine J., Zhang Q., Amonsin A., Alt D., Kapur V.,
 RT Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AE017231; ANS03563.1; -; Genomic DNA.
 DR HSSP; P56690; 1U2S.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0004822; F:isoleucine-tRNA ligase activity; IEA.
 DR GO; GO:0016874; F:Ligase activity; IEA.
 DR GO; GO:0006428; P:isoleucyl-tRNA aminoacylation; IEA.
 DR GO; GO:0006412; P:protein biosynthesis; IEA.
 DR InterPro; IPR002300; tRNA-synt_1a.
 DR InterPro; IPR001412; tRNA-synt_1.
 DR InterPro; IPR002301; tRNA-synt_1le.
 DR Pfam; PF00133; tRNA-synt_1; 1.
 DR PRINTS; PR00984; TRNASYNTHLE.
 DR TIGRFAm; TIGR00392; 11es; 1.
 DR PROSITE; PS00178; AA_trna_ligase_1; UNKNOWN_1.
 KW Complete proteome.
 SO SEQUENCE 1053 AA; 118304 MW; 0F210268429438B0 CRC64;

RESULT 3

Q740U6 MYCBA PRELIMINARY; PRT; 1053 AA.

ID Q740U6 MYCBA PRELIMINARY; PRT; 1053 AA.
 AC Q740U6;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DR Isoeucyl-tRNA synthetase 11es (EC 6.1.1.5).
 GN Name=11es; OrderedlocusNames=MAP1246;
 OS Mycobacterium paratuberculosis.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacteriaceae; Mycobacterium;
 OC Mycobacterium avium complex (MAC).
 OX NCBI_TaxID=1770;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=R10;
 RA Li L., Bannantine J., Zhang Q., Amonsin A., Alt D., Kapur V.,
 RT Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AE017231; ANS03563.1; -; Genomic DNA.
 DR HSSP; P56690; 1U2S.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0004822; F:isoleucine-tRNA ligase activity; IEA.
 DR GO; GO:0016874; F:Ligase activity; IEA.
 DR GO; GO:0006428; P:isoleucyl-tRNA aminoacylation; IEA.
 DR GO; GO:0006412; P:protein biosynthesis; IEA.
 DR InterPro; IPR002300; tRNA-synt_1a.
 DR InterPro; IPR001412; tRNA-synt_1.
 DR InterPro; IPR002301; tRNA-synt_1le.
 DR Pfam; PF00133; tRNA-synt_1; 1.
 DR PRINTS; PR00984; TRNASYNTHLE.
 DR TIGRFAm; TIGR00392; 11es; 1.
 DR PROSITE; PS00178; AA_trna_ligase_1; UNKNOWN_1.
 KW Complete proteome.
 SO SEQUENCE 1053 AA; 118304 MW; 0F210268429438B0 CRC64;

Query Match 84.4%; Score 454; DB 2; Length 1053;
 Best Local Similarity 86.9%; Pred. No. 2.3e-32;
 Matches 93; Conservative 4; Mismatches 10; Indels 0; Gaps 0;

QY 1 MAADPESTALPDGAGLVLDGVTATLEAEAGAKXRIEQLRKSTGLDVSDRIRVVM 60
 ID Q740U6 MYCBA PRELIMINARY; PRT; 1059 AA.
 AC Q740U6;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 13-SEP-2005 (Rel. 48, Last annotation update)
 DR Isoeucyl-tRNA synthetase 11es (EC 6.1.1.5) (Isoleucine--tRNA ligase)
 SO SEQUENCE 1059 AA; 118304 MW; 0F210268429438B0 CRC64;

RESULT 4

Q740U6 MYCBA PRELIMINARY; PRT; 1059 AA.

ID Q740U6 MYCBA PRELIMINARY; PRT; 1059 AA.
 AC Q740U6;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 13-SEP-2005 (Rel. 48, Last annotation update)
 DR Isoeucyl-tRNA synthetase 11es (EC 6.1.1.5) (Isoleucine--tRNA ligase)
 SO SEQUENCE 1059 AA; 118304 MW; 0F210268429438B0 CRC64;

```

DE (11ens).
GN Name=11es; OrderedLocNames=ML1195; ORFNames=MLCB458.10;
OS Mycobacterium leprae.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1769;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=TN;
RX MEDLINE=2118732; PubMed=11234002; DOI=10.1038/35059006;
RA Cole S.T., Eiglmeyer K., Parkhill J., James K.D., Thomson N.R.,
RA Wheeler P.R., Honore N., Garner T., Churcher C.M., Harris D.B.,
RA Whittall K.L., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R.M., Devlin K., Duthoy S., Felwell T., Fraser A., Hamlin N.,
RA Holroyd S., Hornsby T., Jagels K., Lacroix C., Maclean J., Moule S.,
RA Murphy L.D., Oliver K., Quail M.A., Rajadaram M.A., Rutherford K.M.,
RA Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R.,
RA Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,
RA Barrell B.G.;
RT Massive gene decay in the leprosy bacillus.;
RL Nature 409:1007-1011(2001).
CC -1- CATALYTIC ACTIVITY: ATP + L-isoleucine + tRNA(Ile) = AMP +
CC diphosphate + L-isoleucyl-tRNA(Ile).
CC -1- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: Belongs to the class-I aminoacyl-tRNA synthetase
CC family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
DR EMBL; AL049478; CAB39575.1; -; Genomic DNA.
DR EMBL; AL583921; CAC31576.1; -; Genomic DNA.
DR PIR; E87058; E87058.
DR HSP; P56690; 11LE.
DR Leproma; ML1195; -.
DR InterPro; IPR002300; tRNA-synt_1a.
DR InterPro; IPR001412; tRNA-synt_1.
DR InterPro; IPR002301; tRNA-synt_1le.
DR Pfam; PF00133; tRNA-synt_1; 1.
DR PRINTS; PR00984; TRNASYNTHILE.
DR TIGRfam; TIGR00392; 11es; 1.
DR PROSITE; PS00178; AA_TRNA_LIGASE_I; 1.
KW Aminoacyl-tRNA synthetase; ATP-binding; Complete proteome; Ligase;
KW Metal-binding; Nucleotide-binding; Protein biosynthesis; Zinc.
FT MOTIF 59 69 "HIGH" region.
FT MOTIF 637 641 "KMSKS" region.
FT BINDING 640 640 ATP (By similarity).
SQ SEQUENCE 1059 AA; 119809 MW; 67FC7659B399E39 CRC64;

Query Match 78.6%; Score 423; DB 1; Length 1059;
Best Local Similarity 80.4%; Pred. No. 1.5e-29;
Matches 86; Conservative 6; Mismatches 15; Indels 0; Gaps 0;

QY 1 MAADPSTALPGAGIVLDGVTYTALEBAGAKRIRLOLRKSTGIDVSDRIRVYM 60
DB 952 VVAPNFTALPGSSGIVLDVDTYTBELBAGAKRIRLOLRKSTGIDVSDRIRVLM 1011
DB 61 SVPAERDPAKTRDILAGIILATDPEFADLADGVAIGDGVASIEK 107
DB 1012 SVPAERDPAKTRDILAGIILATDPEFADLADGVAIGDGVASIEK 1058

RESULT 5
OS YTM9_NOCFA PRELIMINARY; PRT; 1042 AA.
ID O5YTM9_NOCFA PRELIMINARY; PRT; 1042 AA.
AC O5YTM9;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)

```

```

DE Putative isoleucyl-tRNA synthetase.
GN Name=11es; OrderedLocNames=nfal17760;
OS Nocardia farcinica.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Nocardiaceae; Nocardia.
OX NCBI_TaxID=37329;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=IFM 10152;
RX PubMed=15466710; DOI=10.1073/pnas.0406410101;
RA Ishikawa J., Yamashita A., Mikami Y., Hoshino Y., Kurita H., Hotta K.,
RA Shiba T., Hattori M.;
RT The complete genomic sequence of Nocardia farcinica IFM 10152.;
RL Proc. Natl. Acad. Sci. U.S.A. 101:14925-14930(2004).
CC -1- CATALYTIC ACTIVITY: ATP + L-isoleucine + tRNA(Ile) = AMP +
CC diphosphate + L-isoleucyl-tRNA(Ile).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
DR EMBL; AP006518; BAD56622.1; -; Genomic DNA.
DR GO; GO:0005524; P:ATP binding; IEA.
DR GO; GO:0004822; P:isoleucine-tRNA ligase activity; IEA.
DR GO; GO:0016874; P:ligase activity; IEA.
DR GO; GO:0046872; P:metal ion binding; IEA.
DR GO; GO:0004428; P:isoleucyl-tRNA aminoacylation; IEA.
DR GO; GO:0004412; P:protein biosynthesis; IEA.
DR InterPro; IPR002300; tRNA-synt_1a.
DR InterPro; IPR001412; tRNA-synt_1.
DR InterPro; IPR002301; tRNA-synt_1le.
DR Pfam; PF00133; tRNA-synt_1; 1.
DR PRINTS; PR00984; TRNASYNTHILE.
DR TIGRfam; TIGR00392; 11es; 1.
DR PROSITE; PS00178; AA_TRNA_LIGASE_I; UNKNOWN 1.
KW ATP-binding; Aminoacyl-tRNA synthetase; Complete proteome; Ligase;
KW Metal-binding; Nucleotide-binding; Protein biosynthesis; Zinc.
SQ SEQUENCE 1042 AA; 116567 MW; A0BBB64BBD448E8 CRC64;

Query Match 65.5%; Score 352.5; DB 2; Length 1042;
Best Local Similarity 66.7%; Pred. No. 3.5e-23;
Matches 72; Conservative 9; Mismatches 26; Indels 1; Gaps 1;

QY 1 MAADPSTALPGAGIVLDGVTYTALEBAGAKRIRLOLRKSTGIDVSDRIRVYM 60
DB 934 VVAPNFTALPGAGIVLDVDTYTBELBAGAKRIRLOLRKSTGIDVSDRIRVLM 993
DB 61 SVPAERDPAKTRDILAGIILATDPEFADLADGVAIGDGVASIEK 107
DB 994 EYPAERDPAKTRDILAGIILATDPEFADLADGVAIGDGVASIEK 1041

RESULT 6
OS YTM9_NOCFA PRELIMINARY; PRT; 1103 AA.
ID O8FNV0_COREF PRELIMINARY; PRT; 1103 AA.
AC O8FNV0;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Putative isoleucyl-tRNA synthetase.
GN OrderedLocNames=CE2043;
OS Corynebacterium efficiens.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.
OX NCBI_TaxID=152794;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=YS-314 / AT 12310 / DSM 44549 / JCM 11189;
RX MEDLINE=22723752; PubMed=12840036; DOI=10.1101/gr.1285603;
RA Nishio Y., Nakamura Y., Kawarabayashi Y., Usuda Y., Kimura E.,
RA Sugimoto S., Matsui K., Yamagishi A., Kikuchi H., Ikeo K.,
RA Gojobori T.;
RT "Comparative complete genome sequence analysis of the amino acid
RT replacements responsible for the thermostability of Corynebacterium
RT efficiens.";
RL Genome Res. 13:1572-1579(2003).
DR EMBL; BA000035; BAC18853.1; -; Genomic DNA.

```



```

Db      1070 AKVEQFRLIAHETLATSFE---VKEGAEIG 1097

RESULT 12
ID      Q9RUP8_DEIRA PRELIMINARY; PRT; 1078 AA.
AC      Q9RUP8;
DT      01-MAY-2000 (TrEMBLrel. 13, Created)
DT      01-MAY-2000 (TrEMBLrel. 13, last sequence update)
DE      01-MAR-2004 (TrEMBLrel. 26, last annotation update)
DE      Isolucyl1-cRNA synthetase.
GN      OrderedLocNames=DR1335;
OS      Deinococcus radiodurans.
OC      Bacteria; Deinococcus-Thermus; Deinococci; Deinococcales;
OC      Deinococcaceae; Deinococcus.
OX      NCBI_TaxID=1299;
RN      [1]
RP      NUCLEOTIDE SEQUENCE.
RC      STRAIN=R1 / ATCC 13939 / DSM 20539 / NCIB 9279;
RX      MEDLINE=20036896; PubMed=10567266; DOI=10.1126/science.286.5444.1571;
       White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
       Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
       Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
       Vamathevan J.J., Lam P., McDonald L.A., Utterback T.R., Zalewski C.,
       Vamathevan K.S., Aravind L., Daly M.J., Minton K.W., Fleischman R.D.,
       Ketchum K.A., Nelson K.E., Salzberg S.L., Smith H.O., Venter J.C.,
       Frazer C.M.;
RA      "Genome sequence of the radioresistant bacterium Deinococcus
RT      radiodurans R1.";
RT      Science 286:1571-1577(1999).
DR      EMBL; AE001980; AAF10907.1; -; Genomic_DNA.
DR      PIR; E75407; E75407.
DR      HSSP; P56690; 1ILR.
DR      TIGR; DR1335; -;
DR      GO; GO:0005524; P:ATP binding; IEA.
DR      GO; GO:0004822; P:isoleucine-cRNA ligase activity; IEA.
DR      GO; GO:0006428; P:isoleucyl1-cRNA aminoacylation; IEA.
DR      InterPro; IPR002301; tRNA-synt_1a.
DR      InterPro; IPR002301; tRNA-synt_1le.
DR      Pfam; PF00133; tRNA-synt_1; 1.
DR      PRINTS; PR00984; TRNASYNTHILE.
DR      TIGRPFAM; TIGR00392; 1leS; 1.
DR      Aminoacyl1-cRNA synthetase; Complete proteome.
DR      KW      Aminoacyl1-cRNA synthetase; Complete proteome.
SQ      SEQUENCE 1078 AA; 120272 MW; DF007ED70528F70F CRC64;

Query Match      32.4%; Score 174.5; DB 2; Length 1078;
Best Local Similarity 37.1%; Pred. No. 4.9e-07;
Matches 39; Conservative 16; Mismatches 43; Indels 7; Gaps 1;

QY      2 AADPESTALPDAGIVLVDTVTALAEAGMAKDRIRLEQLRKSTGLDVSRRIVMS 61
DB      972 AKRPEVAAADGIVAVFTALTPELVRGLADLVRAIGERKKAAGFVQRRALALE 1031
QY      62 VPARERDWARTRHDLIAGEIILATDFEPADLADGVALGCVRSIE 106
DB      1032 LDGEALEAAQAMQDFLAGEVLAQVAYGS-----GEGGRAVE 1069

RESULT 13
ID      Q8TN62_METAC PRELIMINARY; PRT; 1058 AA.
AC      Q8TN62;
DT      01-JUN-2002 (TrEMBLrel. 21, Created)
DT      01-JUN-2002 (TrEMBLrel. 21, last sequence update)
DE      01-MAR-2004 (TrEMBLrel. 26, last annotation update)
DE      Isolucyl1-cRNA synthetase.
GN      Name=1leS; OrderedLocNames=MA2431;
OS      Methanosarcina acetivorans.
OC      Archaea; Euryarchaeota; Methanomicrobota; Methanosarcinales;
OC      Methanosarcinaceae; Methanosarcina.
OX      NCBI_TaxID=2214;
RN      [1]
RP      NUCLEOTIDE SEQUENCE.

```

```

RC      STRAIN=C2A / ATCC 35395 / DSM 2834;
RX      MEDLINE=21929760; PubMed=11932238; DOI=10.1101/gr.223902;
RA      Galagan J.E., Nusbaum C., Roy A., Endrizzi M.G., Macdonald P.,
RA      Fitzhugh W., Calvo S., Engels R., Smirnov S., Anoor D., Brown A.,
RA      Allen N., Naylor J., Stange-Thomann N., DeArrelano K., Johnson R.,
RA      Linton L., McEwan P., McKernan K., Talama J., Tirrell A., Ye W.,
RA      Zimmer A., Barber R.D., Cam I., Graham D.E., Graham D.A., Guss A.M.,
RA      Hedderich R., Ingram-Smith C., Kuetner H.C., Krzycki J.A.,
RA      Leigh J.A., Li W., Liu J., Mukhopadhyay B., Reeve J.N., Smith K.,
RA      Springer T.A., Umayam L.A., White O., White R.H., de Macario B.C.,
RA      Perry J.G., Jarrell K.F., Jing H., Macario A.J.L., Paulsen I.T.,
RA      Pritchett M., Sowers K.R., Swanson R.V., Zinder S.H., Lander E.,
RA      Metcalf W.W., Birren B.;
RT      "The genome of Methanosarcina acetivorans reveals extensive metabolic
RT      and physiological diversity.";
RL      Genome Res. 12:532-542(2002).
DR      EMBL; AE010935; AAM05817.1; -; Genomic_DNA.
DR      HSSP; P56690; 1ILR.
DR      GO; GO:0005524; P:ATP binding; IEA.
DR      GO; GO:0004822; P:isoleucine-cRNA ligase activity; IEA.
DR      GO; GO:0016874; P:ligase activity; IEA.
DR      GO; GO:0006428; P:isoleucyl1-cRNA aminoacylation; IEA.
DR      GO; GO:0006412; P:protein biosynthesis; IEA.
DR      InterPro; IPR002300; tRNA-synt_1a.
DR      InterPro; IPR001412; tRNA-synt_1.
DR      InterPro; IPR002301; tRNA-synt_1le.
DR      Pfam; PF00133; tRNA-synt_1; 1.
DR      PRINTS; PR00984; TRNASYNTHILE.
DR      TIGRPFAM; TIGR00392; 1leS; 1.
DR      PROSITE; PS00178; AA_TRNA_LIGASE_I; UNKNOWN_1.
DR      KW      Complete proteome.
SQ      SEQUENCE 1058 AA; 120216 MW; 3940FED68C23A8A3 CRC64;

Query Match      31.0%; Score 167; DB 2; Length 1058;
Best Local Similarity 40.6%; Pred. No. 2.3e-06;
Matches 41; Conservative 17; Mismatches 33; Indels 10; Gaps 2;

QY      5 PESTALPDAGIVLVDTVTALAEAGMAKDRIRLEQLRKSTGLDVSRRIVMSVPA 64
DB      946 PECTAASBDAGIVLVDTVTALAEAGVAREVIRLQMRKELDVVDENIRVSRIEA 1005
QY      65 ERS-DWARTRHDLIAGEIILATDFEPADLADGVALGCVRS 104
DB      1006 EKVTLVETLKDLIAEVRADVD-----LGSISIEVS 1037

RESULT 14
ID      Q9P9L9_METBA PRELIMINARY; PRT; 1058 AA.
AC      Q9P9L9;
DT      01-OCT-2000 (TrEMBLrel. 15, Created)
DT      01-OCT-2000 (TrEMBLrel. 15, last sequence update)
DE      01-MAR-2004 (TrEMBLrel. 26, last annotation update)
DE      Isolucyl1-cRNA synthetase.
GN      Name=1leS;
OS      Methanosarcina barkeri.
OC      Archaea; Euryarchaeota; Methanomicrobota; Methanosarcinales;
OC      Methanosarcinaceae; Methanosarcina.
OX      NCBI_TaxID=2208;
RN      [1]
RP      NUCLEOTIDE SEQUENCE.
RC      STRAIN=Pubaro;
RX      MEDLINE=20225863; PubMed=10762266;
       DOI=10.1128/JB.182.9.2611-2618.2000;
RA      Boccazzzi P., Zhang J.K., Metcalf W.W.;
RT      "Generation of dominant selectable markers for resistance to
RT      pseudomonadic acid by cloning and mutagenesis of the ilS gene from the
RT      archaeon Methanosarcina barkeri furoara.";
RL      J. Bacteriol. 182:2611-2618(2000).
DR      EMBL; AF208389; AAP65673.1; -; Genomic_DNA.
DR      HSSP; P56690; 1ILR.
DR      GO; GO:0005524; P:ATP binding; IEA.
DR      GO; GO:0004822; P:isoleucine-cRNA ligase activity; IEA.

```


This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using SW model

Run on: April 14, 2006, 17:32:59 ; Search time 9.98647 Seconds
(without alignments)
894.107 Million cell updates/sec

Title: US-10-620-246-4

Perfect score: 538
Sequence: 1 MAADPESTALPDAGLVVLT.....ADLADGVAIGDGVRSIEKT 108

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA:*
1: /cgn2_6/ptodata/1/1aa/5_COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/6_COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/H_COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/PCYUS_COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/RE_COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/backfiles.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	538	100.0	108	2	US-09-050-739-4 Sequence 4, Appli
2	485	90.1	1045	1	US-08-452-083-2 Sequence 2, Appli
3	105	19.5	1041	2	US-08-898-978-2 Sequence 2, Appli
4	105	19.5	1041	2	US-09-372-858-2 Sequence 2, Appli
5	99	18.4	1088	1	US-08-742-026-2 Sequence 2, Appli
6	99	18.4	1088	1	US-08-742-026-2 Sequence 2, Appli
7	96	17.8	1072	2	US-09-357-251-31 Sequence 31, Appli
8	86	16.0	1232	2	US-09-583-110-3245 Sequence 3245, Ap
9	78	14.5	209	2	US-08-415-593-43 Sequence 43, Appli
10	78	14.5	978	1	US-08-417-174-27 Sequence 27, Appli
11	76.5	14.2	661	1	US-08-417-174-27 Sequence 121, App
12	76.5	14.2	661	1	US-08-231-565A-27 Sequence 27, Appli
13	76.5	14.2	661	1	US-09-007-961-27 Sequence 27, Appli
14	76.5	14.2	661	1	US-09-267-439-27 Sequence 121, App
15	76.5	14.2	661	2	US-09-267-439-121 Sequence 121, App
16	76.5	14.2	661	2	US-08-388-852B-2 Sequence 27, Appli
17	76.5	14.2	661	2	US-09-073-138-27 Sequence 121, App
18	76.5	14.2	661	2	US-09-073-138-27 Sequence 121, App
19	76.5	14.2	661	2	US-09-073-138-121 Sequence 6, Appli
20	76.5	14.2	661	2	US-09-862-260A-2 Sequence 6, Appli
21	76.5	14.2	661	2	US-07-891-942G-6 Sequence 9821, Ap
22	76	14.1	1054	2	US-09-949-016-9821 Sequence 9821, Ap
23	76	14.1	1054	2	US-09-949-016-9822 Sequence 33, Appli
24	76	14.1	1262	2	US-09-357-251-33 Sequence 6182, Ap
25	76	14.1	1262	2	US-09-949-016-6182 Sequence 6850, Ap
26	76	14.1	1262	2	US-09-949-016-6850 Sequence 4, Appli
27	76	14.1	1266	1	US-08-468-557-4

28	76	14.1	1266	2	US-09-357-251-32	Sequence 32, Appli
29	74.5	13.8	327	1	US-08-463-092B-9	Sequence 9, Appli
30	74.5	13.8	327	1	US-08-460-907B-9	Sequence 9, Appli
31	73.5	13.7	396	2	US-09-712-363-231	Sequence 231, App
32	73.5	13.7	463	2	US-09-902-540-16722	Sequence 23527, A
33	72.5	13.5	311	2	US-09-252-991A-23527	Sequence 22, Appli
34	71.5	13.3	448	2	US-09-878-766A-22	Sequence 8029, Ap
35	71.5	13.3	448	2	US-10-650-369-22	Sequence 26210, A
36	71	13.2	460	2	US-09-252-991A-26210	Sequence 2522, Ap
37	71	13.2	608	2	US-09-605-703B-2522	Sequence 19002, A
38	70.5	13.1	507	2	US-09-231-899-70	Sequence 14084, A
39	70.5	13.1	1481	2	US-09-252-991A-19002	Sequence 5673, Ap
40	70	13.0	462	2	US-09-902-540-14084	Sequence 10017, A
41	70	13.0	966	2	US-09-107-532A-5673	Sequence 24495, A
42	69	12.8	405	2	US-09-949-016-10017	Sequence 6
43	69	12.8	530	2	US-09-252-991A-24495	
44	68.5	12.7	1044	2	US-09-036-987A-6	
45	68.5	12.7	5588	2	US-09-036-987A-6	

ALIGNMENTS

```
RESULT 1
US-09-050-739-4
; Sequence 4, Application US/09050739
; Patent No. 6641814
; GENERAL INFORMATION:
; APPLICANT: ANDERSEN, Peter
; APPLICANT: NIELSEN, Rikke
; APPLICANT: OETTINGER, Thomas
; APPLICANT: RASMUSSEN, Peter Birk
; APPLICANT: ROSENKRANDS, Ida
; APPLICANT: WEIDINGH, Karin
; APPLICANT: FLORIO, Walter
; TITLE OF INVENTION: NUCLEIC ACIDS FRAGMENTS AND POLYPEPTIDE FRAGMENTS
; FILE REFERENCE: 670001-2002.1
; CURRENT APPLICATION NUMBER: US/09/050,739
; CURRENT FILING DATE: 1998-03-30
; EARLIER APPLICATION NUMBER: 0376/97
; EARLIER FILING DATE: 1997-04-02
; EARLIER APPLICATION NUMBER: 1277/97
; EARLIER FILING DATE: 1997-11-10
; EARLIER APPLICATION NUMBER: 60/044,624
; EARLIER FILING DATE: 1997-04-18
; EARLIER APPLICATION NUMBER: 60/070,488
; EARLIER FILING DATE: 1998-01-05
; NUMBER OF SEQ ID NOS: 173
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 108
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-09-050-739-4

Query Match      100.0%; Score 538; DB 2; Length 108;
Best Local Similarity 100.0%; Pred. No. 6.1e-59; Indels 0; Gaps 0;
Matches 108; Conservative 0; Mismatches 0;

QY 1 MAADPESTALPDAGLVVLTGTATLEAGNAKRIRELOEIRKSTGLDVSRRIVM 60
    |||
Db 1 MAADPESTALPDAGLVVLTGTATLEAGNAKRIRELOEIRKSTGLDVSRRIVM 60
    |||
QY 61 SVPARBDMATRRDLIAGRIATDFFPADLADGVAIGDGVRSIEKT 108
    |||
Db 61 SVPARBDMATRRDLIAGRIATDFFPADLADGVAIGDGVRSIEKT 108
    |||

RESULT 2
US-08-452-083-2
; Sequence 2, Application US/08452083
; Patent No. 5756327
```

```

; GENERAL INFORMATION:
; APPLICANT: Sassanfar, Mandana
; TITLE OF INVENTION: RECOMBINANT MYCOBACTERIAL ISOLEUCYL-tRNA
; TITLE OF INVENTION: SYNTHETASE GENES, TESTER STRAINS AND ASSAYS
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Miltia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/452,083
; FILING DATE: 26-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/305,765
; FILING DATE: 13-SEP-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Brook, David E.
; REGISTRATION NUMBER: 22,592
; REFERENCE/DOCKET NUMBER: C9194-08B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 861-6240
; TELEFAX: (617) 861-9540
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1045 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-452-083-2

Query Match          90.1%; Score 485; DB 1; Length 1045;
Best Local Similarity 98.0%; Pred. NO. 5.7e-51;
Matches 97; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAAPSTALPDGAGLVLDGTVTALEAGMAKRIRELOELRKSTGGLDVSRIKRVVM 60
   |||||
DB 944 VAAPSTALPDGAGLVLDGTVTALEAGMAKRIRELOELRKSTGGLDVSRIKRVVM 1003
   |||||

QY 61 SVPAERDWARTRDLIAGEIILATDFEPADLADGVAIGD 99
   |||||
DB 1004 SVPAERDWARTRDLIAGEIILATDFEPADLADGVAIGD 1042
   |||||

RESULT 3
US-08-898-978-2
; Sequence 2, Application US/08898978
; Patent No. 6001602
; GENERAL INFORMATION:
; APPLICANT: Reichard, Raymond W.
; APPLICANT: Brown, James R.
; APPLICANT: Lawlor, Elizabeth J.
; TITLE OF INVENTION: No. 6001602el 1les
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dechert Price & Rhoads
; STREET: 997 Lenox Drive, Building 3, Suite 210
; CITY: Lawrenceville
; STATE: NJ
; COUNTRY: USA
; ZIP: 08543
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
```

```

; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for windows version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/898,978
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Bloom, Allen
; REGISTRATION NUMBER: 29,135
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 609-520-3214
; TELEFAX: 609-520-3259
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1041 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-898-978-2

Query Match          19.5%; Score 105; DB 2; Length 1041;
Best Local Similarity 28.3%; Pred. NO. 0.00074;
Matches 26; Conservative 18; Mismatches 48; Indels 0; Gaps 0;

QY 3 ADPESTRALPDGAGLVLDGTVTALEAGMAKRIRELOELRKSTGGLDVSRIKRVMSV 62
   |||||
DB 931 ASAEGFVARSASFVALDQQLTEPLIMEIASELVNKINTMRNRGLHVSRIALRLHA 990
   |||||

QY 63 PAEREDWARTRDLIAGEIILATDFEPADLADG 94
   |||||
DB 991 PVIQERAFALHKEYICEITLTTSVSVIDYREG 1022
   |||||

RESULT 4
US-09-372-858-2
; Sequence 2, Application US/09372858
; Patent No. 6361970
; GENERAL INFORMATION:
; APPLICANT: Reichard, Raymond
; APPLICANT: Brown, James
; APPLICANT: Lawlor, Elizabeth
; TITLE OF INVENTION: NOVEL 1les
; FILE REFERENCE: CM10051-D1
; CURRENT APPLICATION NUMBER: US/09/372,858
; CURRENT FILING DATE: 1999-12-08
; PRIOR APPLICATION NUMBER: 08/898,978
; PRIOR FILING DATE: 1997-07-23
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1041
; TYPE: PRT
; ORGANISM: Chlamydia trachomatis
; US-09-372-858-2

Query Match          19.5%; Score 105; DB 2; Length 1041;
Best Local Similarity 28.3%; Pred. NO. 0.00074;
Matches 26; Conservative 18; Mismatches 48; Indels 0; Gaps 0;

QY 3 ADPESTRALPDGAGLVLDGTVTALEAGMAKRIRELOELRKSTGGLDVSRIKRVMSV 62
   |||||
DB 931 ASAEGFVARSASFVALDQQLTEPLIMEIASELVNKINTMRNRGLHVSRIALRLHA 990
   |||||

QY 63 PAEREDWARTRDLIAGEIILATDFEPADLADG 94
   |||||
DB 991 PVIQERAFALHKEYICEITLTTSVSVIDYREG 1022
   |||||
```

```
RESULT 5
US-08-742-026-2
; Sequence 2, Application US/08742026
; Patent No. 5885815
; GENERAL INFORMATION:
; APPLICANT: Sasanfar, Mandana
; APPLICANT: Kaufmann, Christoph
; APPLICANT: Gallant, Paul L.
; APPLICANT: Kranz, Janice E.
; APPLICANT: Homan, Fariba
; TITLE OF INVENTION: Candida Isocetyl-cRNA Synthetase
; TITLE OF INVENTION: Proteins, Nucleic Acids and Strains
; TITLE OF INVENTION: Comprising Same
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/742,026
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Brook, David E.
; REGISTRATION NUMBER: 22,592
; REFERENCE/DOCKET NUMBER: CPI95-01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-861-9540
; TELEFAX: 617-861-6240
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1088 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-742-026-2

Query Match      18.4%; Score 99; DB 1; Length 1088;
Best Local Similarity 35.6%; Pred. No. 0.0044;
Matches 21; Conservative 14; Mismatches 24; Indels 0; Gaps 0;

QY      17 LVVLDGTVTALEAGWAKDRIEQLRKSTGLDVSDRIRVWMSVPAEREDWARTRD 75
      ||::||:||||:||||:||||:||||:||||:||||:||||:||||:||||:|
Db      985 LIILDVNLHPLESEGLARLIRIQRLKKAGLNTDDVQVYRVVKDTIDLPKVKD 1043

RESULT 6
US-08-742-026-23
; Sequence 23, Application US/08742026
; Patent No. 5885815
; GENERAL INFORMATION:
; APPLICANT: Sasanfar, Mandana
; APPLICANT: Kaufmann, Christoph
; APPLICANT: Gallant, Paul L.
; APPLICANT: Kranz, Janice E.
; APPLICANT: Homan, Fariba
; TITLE OF INVENTION: Candida Isocetyl-cRNA Synthetase
; TITLE OF INVENTION: Proteins, Nucleic Acids and Strains
; TITLE OF INVENTION: Comprising Same
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
```

```
      COUNTRY: USA
      ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/742,026
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Brook, David E.
; REGISTRATION NUMBER: 22,592
; REFERENCE/DOCKET NUMBER: CPI95-01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-861-9540
; TELEFAX: 617-861-6240
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1088 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-742-026-23

Query Match      18.4%; Score 99; DB 1; Length 1088;
Best Local Similarity 35.6%; Pred. No. 0.0044;
Matches 21; Conservative 14; Mismatches 24; Indels 0; Gaps 0;

QY      17 LVVLDGTVTALEAGWAKDRIEQLRKSTGLDVSDRIRVWMSVPAEREDWARTRD 75
      ||::||:||||:||||:||||:||||:||||:||||:||||:||||:||||:|
Db      985 LIILDVNLHPLESEGLARLIRIQRLKKAGLNTDDVQVYRVVKDTIDLPKVKD 1043

RESULT 7
US-09-357-251-31
; Sequence 31, Application US/09357251
; Patent No. 6271441
; GENERAL INFORMATION:
; APPLICANT: Falco, S. Carl
; APPLICANT: Famodu, Layo O.
; APPLICANT: Orozco, Buddy
; APPLICANT: Schwaber, James S.
; TITLE OF INVENTION: Plant Aminoacyl-cRNA Synthetase
; FILE REFERENCE: BB-1193
; CURRENT APPLICATION NUMBER: US/09/357,251
; CURRENT FILING DATE: 1999-07-20
; EARLIER APPLICATION NUMBER: 60/093,530
; EARLIER FILING DATE: July 21, 1998
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 31
; LENGTH: 1072
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
; US-09-357-251-31

Query Match      17.8%; Score 96; DB 2; Length 1072;
Best Local Similarity 26.7%; Pred. No. 0.01;
Matches 23; Conservative 25; Mismatches 34; Indels 4; Gaps 2;

QY      17 LVVLDGTVTALEAGWAKDRIEQLRKSTGLDVSDRIRVWMSVPAEREDW---ARTH 73
      ||::||:||||:||||:||||:||||:||||:||||:||||:||||:||||:|
Db      971 LIIMDTNIVSELKSGLARLVRIQRLKKCGLENTDDVLEVEYLVKDTIDPEATVKEH 1030

QY      74 RDLIAGEIILATPEFADLADGVAIGD 99
      ||::||:||||:||||:||||:||||:||||:||||:||||:||||:||||:|
Db      1031 FDLWS-KTGRSDIAKYDGSKTPIGD 1055

RESULT 8
US-09-302-540-14264
```

```
/ Sequence 14264, Application US/0902540
/ Patent No. 6833447
/ GENERAL INFORMATION:
/ APPLICANT: Goldman, Barry S.
/ APPLICANT: Hinkle, Gregory J.
/ APPLICANT: Slater, Steven C.
/ APPLICANT: Wiegand, Roger C.
/ TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
/ FILE REFERENCE: 38-10(15849)B
/ CURRENT APPLICATION NUMBER: US/09/902,540
/ PRIOR FILING DATE: 2001-07-10
/ PRIOR APPLICATION NUMBER: 60/217,883
/ PRIOR FILING DATE: 2000-07-10
/ NUMBER OF SEQ ID NOS: 16825
/ SEQ ID NO 14264
/ LENGTH: 1232
/ TYPE: PRT
/ ORGANISM: Myxococcus xanthus
US-09-902-540-14264

Query Match          16.0%; Score 86; DB 2; Length 1232;
Best Local Similarity 34.2%; Pred. No. 0.21;
Matches 26; Conservative 13; Mismatches 37; Indels 0; Gaps 0;

QY 9 AALPDGAGLVLDGTVTAELAEAGMAKDRIEQLRKSTGLDVSDRIRVMSVPAERD 68
DB 1127 AAGMGVGVVLTETLESIVDEGLVRELLARVQARKDWELGYTRICQLMVDGDARVKR 1186
QY 69 WARTHRDLTAGEIAT 84
DB 1187 VTDEARELIARETLAS 1202

RESULT 9
US-09-583-110-3245
/ Sequence 3245, Application US/09583110
/ Patent No. 6699703
/ GENERAL INFORMATION:
/ APPLICANT: Lynn Doucette-Stamm et al.
/ TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
/ TITLE OF INVENTION: Pneumoniae for Diagnostics and Therapeutics
/ FILE REFERENCE: PATH00-07A
/ CURRENT APPLICATION NUMBER: US/09/583,110
/ PRIOR FILING DATE: 2000-05-26
/ PRIOR APPLICATION NUMBER: US 09/107,433
/ PRIOR FILING DATE: 1998-06-30
/ PRIOR APPLICATION NUMBER: US 60/085,131
/ PRIOR FILING DATE: 1998-05-12
/ PRIOR APPLICATION NUMBER: US 60/051,553
/ PRIOR FILING DATE: 1997-07-02
/ NUMBER OF SEQ ID NOS: 5322
/ SEQ ID NO 3245
/ LENGTH: 209
/ TYPE: PRT
/ ORGANISM: Streptococcus pneumoniae
US-09-583-110-3245

Query Match          14.5%; Score 78; DB 2; Length 209;
Best Local Similarity 30.2%; Pred. No. 0.17;
Matches 29; Conservative 13; Mismatches 36; Indels 18; Gaps 3;

QY 14 GAGLVLDGTVTAEL-----EAGMAKDRIEQLRKSTGLDVSDRIRVMSVPAE 65
DB 99 GAGSLVIDETKTRLDTSFPPEPFSKLADQADELDLTKNNLNTFVSPAVDFIP-- 156
QY 66 REDWARTHRDLTAGEIATD-----PEPDLADGV 95
DB 157 --DGEKGYIILAGEIFTTEKGISQTSYADVAIGL 190

RESULT 10
US-08-415-593-43
/ Sequence 43, Application US/08415593
```

```
/ Patent No. 5912140
/ Patent No. 5912140 5776726
/ GENERAL INFORMATION:
/ APPLICANT: Whoriskey, Susan K.
/ APPLICANT: Quinn, Cheryl L.
/ APPLICANT: Tao, Nisajun
/ APPLICANT: Politis-Vitk, Karen I.
/ APPLICANT: Schimmel, Paul R.
/ TITLE OF INVENTION: Recombinant Pneumocystis Aminoacyl tRNA
/ NUMBER OF SEQUENCES: 46
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
/ STREET: Two Millitia Drive
/ CITY: Lexington
/ STATE: Massachusetts
/ COUNTRY: U.S.
/ ZIP: 02173
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/415,593
/ FILING DATE:
/ CLASSIFICATION: 435
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Brook, David E.
/ REGISTRATION NUMBER: 22,592
/ REFERENCE/DOCKET NUMBER: CPT94-09
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 617-861-6240
/ TELEFAX: 617-862-9540
/ INFORMATION FOR SEQ ID NO: 43:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 978 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: peptide
US-08-415-593-43

Query Match          14.5%; Score 78; DB 1; Length 978;
Best Local Similarity 32.9%; Pred. No. 1.5;
Matches 24; Conservative 12; Mismatches 33; Indels 4; Gaps 2;

QY 18 VVLDGTVTAELAEAGMAKDRIEQLRKSTGLDVSDRIR---VWMSVPAERDMARTR 74
DB 874 ILLDTKIVPELKTREYIVREVIRVORLRKKVGLQVIDIMEYVITIDSLGLEDAISQHO 933
QY 75 DLAGEIATDFE 87
DB 934 ILLT-KILRRPLE 945

RESULT 11
US-08-417-174-27
/ Sequence 27, Application US/08417174
/ Patent No. 5844075
/ GENERAL INFORMATION:
/ APPLICANT: KANAKAWI, YUTAKA, ROSENBERG,
/ APPLICANT: STEVEN A.
/ TITLE OF INVENTION: MELANOMA ANTIGENS AND
/ TITLE OF INVENTION: THEIR USE IN DIAGNOSTIC AND THERAPEUTIC
/ METHODS
/ NUMBER OF SEQUENCES: 126
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
/ STREET: 345 PARK AVENUE
/ CITY: NEW YORK
/ STATE: NEW YORK
/ COUNTRY: USA
/ ZIP: 10154
```

```

? NAME: CAROL M. GRUPPI
? REGISTRATION NUMBER: 37,341
? REFERENCE/DOCKET NUMBER: 2026-4124US1
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (212) 758-4800
? TELEFAX: (212) 751-6849
? TELERX: 421792
? INFORMATION FOR SEQ ID NO: 121:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 661
? TYPE: amino acid
? STRANDEDNESS: Unknown
? TOPOLOGY: Unknown
? MOLECULE TYPE: Protein
? US-08-417-174-121

Query Match      14.2%; Score 76.5; DB 1; Length 661;
Beet Local Similarity 25.2%; Pred. No. 1.3;
Matches 30; Conservative 16; Mismatches 44; Indels 29; Gaps 5.

Cy    5 PESTPALPDGAGVIVLDGTVAELAEAGMADRIREL-----QELRKST 48
      |||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db    395 PEAAGMTPEAVSYIVLSGTTAAQTTEWEVTARELPDIPPEBGPDAASIMSTBSTISL 454
      :|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Oy    49 G--LDVSDRIRVV--MSVPAS----REDWARTHRDLIGEILATDFEFDALDGVAIGD 100
      |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db    455 GPLIDGTNTRLRVKQVPLDCVLRYKPSFTLIDVQS-----IESAKITQAAPSGEG 507

```

RESULT 13
US-08-231-565A-27
; Sequence 27, Application US/08231565A

```

1 GENERAL INFORMATION:
2 APPLICANT: KAKAKAMI, YUTAKA; ROSENBERG,
3 APPLICANT: STEVEN A.
4 TITLE OF INVENTION: MELANOMA ANTIGENS AND
5 TITLE OF INVENTION: THEIR USE IN DIAGNOSTIC AND THERAPEUTIC
6 TITLE OF INVENTION: METHODS
7 NUMBER OF SEQUENCES: 43

```

```

1 STREET: 345 PARK AVENUE
2 CITY: NEW YORK
3 STATE: NEW YORK
4 COUNTRY: USA
5 ZIP: 10154
6 COMPUTER READABLE FORM:
7 MEDIUM TYPE: FLOPPY DISK
8 COMPUTER: IBM PC COMPATIBLE
9 OPERATING SYSTEM: PC-DOS/MS-DOS
10 SOFTWARE: ASCII
11 CURRENT APPLICATION DATA:
12 APPLICATION NUMBER: US/08/231,565A
13 FILING DATE: 22-APR-1994
14 CLASSIFICATION: 435
15 ATTORNEY/AGENT INFORMATION:
16 NAME: CAROL M. GRUPPI
17 REGISTRATION NUMBER: 37,341
18 REFERENCE/DOCKET NUMBER: 2026-4124
19 TELECOMMUNICATION INFORMATION:
20 TELEPHONE: (212) 758-4800
21 TELEFAX: (212) 751-6849
22 TELEX: 421792
23 INFORMATION FOR SEQ ID NO: 27:
24 SEQUENCE CHARACTERISTICS:
25 LENGTH: 661
26 TYPE: amino acid
27 STRANDEDNESS: Unknown
28 TOPOLOGY: Unknown
29 MOLECULE TYPE: Protein
30 US-08-231-565A-27

```


GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocelebration Ltd.

OM protein - protein search, using sw model

Run on: April 14, 2006, 18:27:13 ; Search time 37.8999 Seconds
(without alignments)
1190.652 Million cell updates/sec

Title: US-10-620-246-4

Perfect score: 538
Sequence: 1 MAADPESTALPDGAGLVTL.....ADLADGVALGDGVRVSIKKT 108

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 10%
Listing first 45 summaries

Database : Published Applications_AA_Main:
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	538	100.0	108	US-09-791-171-4	Sequence 4, Appl1
2	538	100.0	108	US-09-804-980-4	Sequence 4, Appl1
3	538	100.0	108	US-10-620-246-4	Sequence 4, Appl1
4	534	99.3	1041	US-10-282-122A-62617	Sequence 62617, A
5	534	99.3	1041	US-10-282-122A-64571	Sequence 64571, A
6	449	83.5	1061	US-10-282-122A-61984	Sequence 61984, A
7	423	78.6	1059	US-10-282-122A-63826	Sequence 63826, A
8	266.5	49.5	1052	US-10-282-122A-54011	Sequence 54011, A
9	266.5	49.5	1054	US-09-738-626-5852	Sequence 5852, Ap
10	266.5	49.5	1063	US-10-494-541-76	Sequence 76, Appl
11	147.5	27.4	1047	US-10-156-761-13657	Sequence 13657, A
12	132	24.5	1091	US-10-282-122A-76512	Sequence 76512, A
13	128.5	23.9	1035	US-10-282-122A-51998	Sequence 51998, A
14	105	19.5	1036	US-10-503-185-125	Sequence 125, App
15	98	18.2	1088	US-10-032-585-7198	Sequence 7198, Ap
16	97.5	18.1	1026	US-10-282-122A-55127	Sequence 55127, A
17	97	18.0	1026	US-10-724-972A-7479	Sequence 7479, A
18	97	18.0	1085	US-10-282-122A-48671	Sequence 48671, A
19	96	17.8	1072	US-09-864-464-311	Sequence 31, Appl
20	93	17.3	1042	US-10-282-122A-47286	Sequence 47286, A
21	87	16.2	1038	US-10-282-122A-53635	Sequence 53635, A
22	82.5	15.3	659	US-10-369-493-18619	Sequence 18619, A
23	82.5	15.3	7068	US-10-203-295-20	Sequence 20, Appl
24	82.5	15.3	9477	US-10-203-295-37	Sequence 37, Appl
25	82.5	15.3	11088	US-10-203-295-7	Sequence 7, Appl
26	82.5	15.3	11096	US-10-732-923-20557	Sequence 20557, A
27	79	14.7	375	US-10-617-038-33	Sequence 33, Appl

28	78.5	14.6	329	4	US-10-282-122A-62051	Sequence 62051, A
29	78.5	14.6	1077	4	US-10-128-714-3179	Sequence 3179, Ap
30	78.5	14.6	1077	4	US-10-128-714-8179	Sequence 8179, Ap
31	76.5	14.2	375	4	US-10-156-761-12584	Sequence 12584, A
32	76.5	14.2	575	4	US-10-287-971-46	Sequence 46, Appl
33	76.5	14.2	660	4	US-10-245-871-185	Sequence 385, App
34	76.5	14.2	660	4	US-10-253-286-385	Sequence 385, App
35	76.5	14.2	661	3	US-09-862-260A-2	Sequence 2, Appl1
36	76.5	14.2	661	3	US-09-812-238B-2	Sequence 2, Appl1
37	76.5	14.2	661	3	US-09-898-860-27	Sequence 27, Appl
38	76.5	14.2	661	3	US-09-898-860-121	Sequence 121, App
39	76.5	14.2	661	4	US-10-207-655-77	Sequence 77, Appl
40	76.5	14.2	661	4	US-10-136-145-2	Sequence 2, Appl1
41	76.5	14.2	661	4	US-10-117-937-70	Sequence 70, Appl
42	76.5	14.2	661	4	US-10-168-417A-2	Sequence 2, Appl1
43	76.5	14.2	661	4	US-10-296-734-817	Sequence 817, App
44	76.5	14.2	661	4	US-10-657-022-70	Sequence 70, Appl
45	76.5	14.2	661	4	US-10-685-977-27	Sequence 27, Appl

ALIGNMENTS

```
RESULT 1
US-09-791-171-4
; Sequence 4, Application US/09791171
; Patent No. US20020094336A1
; GENERAL INFORMATION:
; APPLICANT: ANDERSEN, Peter
; APPLICANT: NIELSEN, Rikke
; APPLICANT: OETTINGER, Thomas
; APPLICANT: RASMUSSEN, Peter Blak
; APPLICANT: ROSENKRANDS, Ida
; APPLICANT: WELDINGH, Karin
; APPLICANT: FLORIO, Walter
; TITLE OF INVENTION: NUCLEIC ACIDS FRAGMENTS AND POLYPEPTIDE FRAGMENTS
; FILE REFERENCE: 670001-2002.1
; CURRENT APPLICATION NUMBER: US/09/791,171
; CURRENT FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 09/050,739
; PRIOR FILING DATE: 1998-03-30
; PRIOR APPLICATION NUMBER: 0376/97
; PRIOR FILING DATE: 1997-04-02
; PRIOR APPLICATION NUMBER: 1277/97
; PRIOR FILING DATE: 1997-11-10
; PRIOR APPLICATION NUMBER: 60/044,624
; PRIOR FILING DATE: 1997-04-18
; PRIOR APPLICATION NUMBER: 60/070,488
; PRIOR FILING DATE: 1998-01-05
; NUMBER OF SEQ ID NOS: 173
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 108
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-09-791-171-4
Query Match 100.0%; Score 538; DB 3; Length 108;
Best Local Similarity 100.0%; Pred. No. 6.5e-53;
Matches 108; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MAADPESTALPDGAGLVVDGTVTAELEAGNKRIRREIQELRKSTGLDVSRRIVM 60
Db 1 MAADPESTALPDGAGLVVDGTVTAELEAGNKRIRREIQELRKSTGLDVSRRIVM 60
QY 61 SVPAERDMARTRDILAGIILATDFEFADLADGVALGDGVRVSIKKT 108
Db 61 SVPAERDMARTRDILAGIILATDFEFADLADGVALGDGVRVSIKKT 108
RESULT 2
US-09-804-980-4
```

```
/ Sequence 4, Application US/09804980
/ Publication No. US20030147897A1
/ GENERAL INFORMATION:
/ APPLICANT: Statens Serum Institut
/ APPLICANT: Andersen, Peter
/ TITLE OF INVENTION: M. Tuberculosis Antigena
/ FILE REFERENCE: 670001-2002.4
/ CURRENT APPLICATION NUMBER: US/09/804,980
/ CURRENT FILING DATE: 2001-03-12
/ NUMBER OF SEQ ID NOS: 257
/ SOFTWARE: PatentIn version 3.0
/ SEQ ID NO 4
/ LENGTH: 108
/ TYPE: PR1
/ ORGANISM: Mycobacterium tuberculosis
US-09-804-980-4

Query Match      100.0%; Score 538; DB 3; Length 108;
Best Local Similarity 100.0%; Pred. No. 6.5e-53;
Matches 108; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MAADPESTALPGAGLVVDTGVTATLEAGNAKORIRELOIRKSTGLDVSDRIRVVM 60
        |||||||
DB      1 MAADPESTALPGAGLVVDTGVTATLEAGNAKORIRELOIRKSTGLDVSDRIRVVM 60

QY      61 SVAEREDWARTHRDLIAGEILATDPEFADLDGVALGQGVRSIEKT 108
        |||||||
DB      61 SVAEREDWARTHRDLIAGEILATDPEFADLDGVALGQGVRSIEKT 108

RESULT 3
US-10-620-246-4
/ Sequence 4, Application US/10620246
/ Publication No. US20040115211A1
/ GENERAL INFORMATION:
/ APPLICANT: ANDERSEN, Peter
/ APPLICANT: NIELSEN, Rikke
/ APPLICANT: OETTINGER, Thomas
/ APPLICANT: RASMUSSEN, Peter Blvk
/ APPLICANT: ROSENBRANDS, Ida
/ APPLICANT: WELDLING, Karin
/ APPLICANT: FLORIO, Walter
/ TITLE OF INVENTION: NUCLEIC ACIDS FRAGMENTS AND POLYPEPTIDE FRAGMENTS
/ FILE REFERENCE: 670001-2002.1A
/ CURRENT APPLICATION NUMBER: US/10/620,246
/ CURRENT FILING DATE: 2003-07-15
/ PRIOR APPLICATION NUMBER: 09/050,739
/ PRIOR FILING DATE: 1998-03-30
/ PRIOR APPLICATION NUMBER: 0376/97
/ PRIOR FILING DATE: 1997-04-02
/ PRIOR APPLICATION NUMBER: 1277/97
/ PRIOR FILING DATE: 1997-11-10
/ PRIOR APPLICATION NUMBER: 60/044,624
/ PRIOR FILING DATE: 1997-04-18
/ PRIOR APPLICATION NUMBER: 60/070,488
/ PRIOR FILING DATE: 1998-01-05
/ PRIOR APPLICATION NUMBER: 10/138,473
/ PRIOR FILING DATE: 2002-05-02
/ PRIOR APPLICATION NUMBER: 09/791,171
/ PRIOR FILING DATE: 2001-02-20
/ PRIOR APPLICATION NUMBER: 09/415,884
/ PRIOR FILING DATE: 1999-10-08
/ PRIOR APPLICATION NUMBER: 60/116,673
/ PRIOR FILING DATE: 1999-01-21
/ PRIOR APPLICATION NUMBER: 1281/98
/ PRIOR FILING DATE: 1998-10-08
/ NUMBER OF SEQ ID NOS: 173
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 4
/ LENGTH: 108
/ TYPE: PR1
/ ORGANISM: Mycobacterium tuberculosis
```

```
US-10-620-246-4

Query Match      100.0%; Score 538; DB 4; Length 108;
Best Local Similarity 100.0%; Pred. No. 6.5e-53;
Matches 108; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MAADPESTALPGAGLVVDTGVTATLEAGNAKORIRELOIRKSTGLDVSDRIRVVM 60
        |||||||
DB      1 MAADPESTALPGAGLVVDTGVTATLEAGNAKORIRELOIRKSTGLDVSDRIRVVM 60

QY      61 SVAEREDWARTHRDLIAGEILATDPEFADLDGVALGQGVRSIEKT 108
        |||||||
DB      61 SVAEREDWARTHRDLIAGEILATDPEFADLDGVALGQGVRSIEKT 108
```

```
RESULT 4
US-10-282-122A-62617
/ Sequence 62617, Application US/10282122A
/ Publication No. US20040029129A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, Liangau
/ APPLICANT: Zamudio, Carlos
/ APPLICANT: Malone, Cheryl
/ APPLICANT: Haselbeck, Robert
/ APPLICANT: Ohlsen, Karl
/ APPLICANT: Zyskind, Judith
/ APPLICANT: Wall, Daniel
/ APPLICANT: Trawick, John
/ APPLICANT: Carr, Grant
/ APPLICANT: Yamamoto, Robert
/ APPLICANT: Forsyth, R.
/ APPLICANT: Xu, H.
/ TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
/ FILE REFERENCE: E1179A.034A
/ CURRENT APPLICATION NUMBER: US/10/282,122A
/ CURRENT FILING DATE: 2003-02-20
/ PRIOR APPLICATION NUMBER: 60/191,078
/ PRIOR FILING DATE: 2000-03-21
/ PRIOR APPLICATION NUMBER: 60/206,848
/ PRIOR FILING DATE: 2000-05-23
/ PRIOR APPLICATION NUMBER: 60/207,727
/ PRIOR FILING DATE: 2000-05-26
/ PRIOR APPLICATION NUMBER: 60/230,335
/ PRIOR FILING DATE: 2000-09-06
/ PRIOR APPLICATION NUMBER: 60/230,347
/ PRIOR FILING DATE: 2000-09-09
/ PRIOR APPLICATION NUMBER: 60/242,578
/ PRIOR FILING DATE: 2000-10-23
/ PRIOR APPLICATION NUMBER: 60/253,625
/ PRIOR FILING DATE: 2000-11-27
/ PRIOR APPLICATION NUMBER: 60/257,931
/ PRIOR FILING DATE: 2000-12-22
/ PRIOR APPLICATION NUMBER: 60/267,636
/ PRIOR FILING DATE: 2001-02-09
/ PRIOR APPLICATION NUMBER: 60/269,308
/ PRIOR FILING DATE: 2001-02-16
/ Remaining Prior Application data removed - See file Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 78614
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 62617
/ LENGTH: 1041
/ TYPE: PR1
/ ORGANISM: Mycobacterium bovis
US-10-282-122A-62617
```

```
Query Match      99.3%; Score 534; DB 4; Length 1041;
Best Local Similarity 99.1%; Pred. No. 3.5e-51;
Matches 107; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 MAADPESTALPGAGLVVDTGVTATLEAGNAKORIRELOIRKSTGLDVSDRIRVVM 60
        :|||||
DB      934 VMAADPESTALPGAGLVVDTGVTATLEAGNAKORIRELOIRKSTGLDVSDRIRVVM 933
```

QY 61 SVPAERBDMWATHRDLIAGEIILATDPFPAADLADGVAIGDGVRSIEXT 108
 Db 994 SVPAERBDMWATHRDLIAGEIILATDPFPAADLADGVAIGDGVRSIEXT 1041

RESULT 5

US-10-282-122A-64571
 ; Sequence 64571, Application US/10282122A
 ; Publication No. US20040029129A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, Liangsu
 ; APPLICANT: Zamudio, Carlos
 ; APPLICANT: Malone, Cheryl
 ; APPLICANT: Haselbeck, Robert
 ; APPLICANT: Ohlsen, Karl
 ; APPLICANT: Zyskind, Judith
 ; APPLICANT: Wall, Daniel
 ; APPLICANT: Trawick, John
 ; APPLICANT: Carr, Grant
 ; APPLICANT: Yamamoto, Robert
 ; APPLICANT: Forsyth, R.
 ; APPLICANT: Xu, H.
 ; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
 ; FILE REFERENCE: ELITRA.034A
 ; CURRENT APPLICATION NUMBER: US/10/282,122A
 ; CURRENT FILING DATE: 2003-02-20
 ; PRIOR APPLICATION NUMBER: 60/191,078
 ; PRIOR FILING DATE: 2000-03-21
 ; PRIOR APPLICATION NUMBER: 60/206,848
 ; PRIOR FILING DATE: 2000-05-23
 ; PRIOR APPLICATION NUMBER: 60/207,727
 ; PRIOR FILING DATE: 2000-05-26
 ; PRIOR APPLICATION NUMBER: 60/230,335
 ; PRIOR FILING DATE: 2000-09-06
 ; PRIOR APPLICATION NUMBER: 60/230,347
 ; PRIOR FILING DATE: 2000-09-09
 ; PRIOR APPLICATION NUMBER: 60/242,578
 ; PRIOR FILING DATE: 2000-10-23
 ; PRIOR APPLICATION NUMBER: 60/253,625
 ; PRIOR FILING DATE: 2000-11-27
 ; PRIOR APPLICATION NUMBER: 60/257,931
 ; PRIOR FILING DATE: 2000-12-22
 ; PRIOR APPLICATION NUMBER: 60/267,636
 ; PRIOR FILING DATE: 2001-02-09
 ; PRIOR APPLICATION NUMBER: 60/269,308
 ; PRIOR FILING DATE: 2001-02-16
 ; Remaining Prior Application data removed - See File Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 78614
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 64571
 ; LENGTH: 1041
 ; TYPE: PRT
 ; ORGANISM: Mycobacterium tuberculosis
 ; US-10-282-122A-64571

Query Match 99.3%; Score 534; DB 4; Length 1041;
 Best Local Similarity 99.1%; Pred. No. 3,5e-51;
 Matches 107; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAADPESTALPDGAGIVVDGVTVAELBAEGNAKDIRIRLOELRKSTGLDVSDRIRVM 60
 Db 934 VAADPESTALPDGAGIVVDGVTVAELBAEGNAKDIRIRLOELRKSTGLDVSDRIRVM 993
 QY 61 SVPAERBDMWATHRDLIAGEIILATDPFPAADLADGVAIGDGVRSIEXT 108
 Db 994 SVPAERBDMWATHRDLIAGEIILATDPFPAADLADGVAIGDGVRSIEXT 1041

RESULT 6

US-10-282-122A-61984
 ; Sequence 61984, Application US/10282122A
 ; Publication No. US20040029129A1
 ; GENERAL INFORMATION:

APPLICANT: Wang, Liangsu
 APPLICANT: Zamudio, Carlos
 APPLICANT: Malone, Cheryl
 APPLICANT: Haselbeck, Robert
 APPLICANT: Ohlsen, Karl
 APPLICANT: Zyskind, Judith
 APPLICANT: Wall, Daniel
 APPLICANT: Trawick, John
 APPLICANT: Carr, Grant
 APPLICANT: Yamamoto, Robert
 APPLICANT: Forsyth, R.
 APPLICANT: Xu, H.
 TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
 FILE REFERENCE: ELITRA.034A
 CURRENT APPLICATION NUMBER: US/10/282,122A
 CURRENT FILING DATE: 2003-02-20
 PRIOR APPLICATION NUMBER: 60/191,078
 PRIOR FILING DATE: 2000-03-21
 PRIOR APPLICATION NUMBER: 60/206,848
 PRIOR FILING DATE: 2000-05-23
 PRIOR APPLICATION NUMBER: 60/207,727
 PRIOR FILING DATE: 2000-05-26
 PRIOR APPLICATION NUMBER: 60/230,335
 PRIOR FILING DATE: 2000-09-06
 PRIOR APPLICATION NUMBER: 60/230,347
 PRIOR FILING DATE: 2000-09-09
 PRIOR APPLICATION NUMBER: 60/242,578
 PRIOR FILING DATE: 2000-10-23
 PRIOR APPLICATION NUMBER: 60/253,625
 PRIOR FILING DATE: 2000-11-27
 PRIOR APPLICATION NUMBER: 60/257,931
 PRIOR FILING DATE: 2000-12-22
 PRIOR APPLICATION NUMBER: 60/267,636
 PRIOR FILING DATE: 2001-02-09
 PRIOR APPLICATION NUMBER: 60/269,308
 PRIOR FILING DATE: 2001-02-16
 Remaining Prior Application data removed - See File Wrapper or PALM.
 NUMBER OF SEQ ID NOS: 78614
 SOFTWARE: PatentIn version 3.1
 SEQ ID NO 61984
 LENGTH: 1061
 TYPE: PRT
 ORGANISM: Mycobacterium avium
 US-10-282-122A-61984

Query Match 83.5%; Score 449; DB 4; Length 1061;
 Best Local Similarity 86.0%; Pred. No. 1.5e-41;
 Matches 92; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

QY 1 MAADPESTALPDGAGIVVDGVTVAELBAEGNAKDIRIRLOELRKSTGLDVSDRIRVM 60
 Db 954 VAADPESTALPDGAGIVVDGVTVAELBAEGNAKDIRIRLOELRKSTGLDVSDRIRVM 1013
 QY 61 SVPAERBDMWATHRDLIAGEIILATDPFPAADLADGVAIGDGVRSIEXT 107
 Db 1014 SVPAERBDMWATHRDLIAGEIILATDPFPAADLADGVAIGDGVRSIEXT 1060

RESULT 7

US-10-282-122A-63826
 ; Sequence 63826, Application US/10282122A
 ; Publication No. US20040029129A1
 ; GENERAL INFORMATION:

APPLICANT: Wang, Liangsu
 APPLICANT: Zamudio, Carlos
 APPLICANT: Malone, Cheryl
 APPLICANT: Haselbeck, Robert
 APPLICANT: Ohlsen, Karl
 APPLICANT: Zyskind, Judith
 APPLICANT: Wall, Daniel
 APPLICANT: Trawick, John
 APPLICANT: Carr, Grant
 APPLICANT: Yamamoto, Robert

! PRIOR FILING DATE: 2000-05-23

Best Local Similarity 53.6

Query Match	49.5%;	Score 266.5;	DB 3;	Length 1054;
Best Local Similarity	53.6%;	Pred. No. 7.5e-21;		

```

; LENGTH: 1047
; TYPE: PRT
; ORGANISM: Streptomyces avermectilis
US-10-156-761-13657

Query Match 27.4%; Score 147.5; DB 4; Length 1047;
Best Local Similarity 37.8%; Pred. No. 2,36-07;
Matches 37; Conservative 11; Mismatches 31; Indels 19; Gaps 2

Qy
6 ESTAALPDGAGLVVLDGTTVALEABGMAQRIEQLRKSGLDVSDRIRVMSVAP 65
Db
939 EGMVSAVSDGATVADLDEITTEELQAGIARDAIRLIGARNSGIDVADRIAL----- 991
Qy
66 REDWART-----HRDLIAEILATDEPFDLAD 93
Db
992 --RWTSTDPVIAALSERHSLIADVATPDAGEAD 1027

RESULT 12
US-10-282-122A-76512
; Sequence 76512, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Habelbeck, Robert
; APPLICANT: Ohlsen, Karl
; APPLICANT: Zykkind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OR INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ERIITA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; PRIOR FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 76512
; LENGTH: 1091
; TYPE: PRT
; ORGANISM: Treponema pallidum
US-10-282-122A-76512

Query Match 24.5%; Score 132; DB 4; Length 1091;
Best Local Similarity 37.2%; Pred. No. 1,4e-05;
Matches 42; Conservative 16; Mismatches 43; Indels 12; Gaps 5

6 ESTAALPDGAGLVVLDGTTVALEABGMAQRIEQLRKSGLDVSDRIRVMS-- 61

```

Db 979 ESJKAINEGTLTALPTTLTLEEDLLBEAIDLVKGVNLRKRGFSLVDRICLRVSSDQ 1038
QY 62 -VPAEREDMARTHDLIAGEIATDPEFADLADGVAI----GQGV--RVSIEK 107
Db 1039 DIVCAKAY-DLHRSTYVGETTLAAHVOMANVRDGAANVYKSDANVMEVSDK 1090

RESULT 13

US-10-282-122A-51998
; Sequence 51998, Application US/10282122A
; Publication No. US20040029128A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See file wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 51998
; LENGTH: 1035
; TYPE: PRT
; ORGANISM: Clostridium acetobutylicum
US-10-282-122A-51998

Query Match 23.9%; Score 128.5; DB 4; Length 1035;
Best Local Similarity 30.1%; Pred. No. 3.2e-05;
Matches 28; Conservative 26; Mismatches 36; Indels 3; Gaps 1;
QY 18 VVLDGTVTALEAEAGMADRIKRELQELKSTGLDVSRRIRVMSVPAEREDMARTHDLI 77
Db 941 IVLETTITDLREBGRVRLSKVQNRKSGFVVALKIELIYSGNEKLENVIRKPEDTI 1000
QY 78 AGEIATDPEFADLADGVAI---GQGVRSIEK 107
Db 1001 KKETLATDIIYSENKEAAYNINGEELNVFVK 1033

RESULT 14
US-10-503-135-125

; Sequence 125, Application US/10503135
; Publication No. US20050152926A1
; GENERAL INFORMATION:
; APPLICANT: BENSI Giuliano
; APPLICANT: GRANDI Guido
; TITLE OF INVENTION: CYTOTOXIC T-CELL EPITOPES FROM CHLAMYDIA
; FILE REFERENCE: 002441.00089
; CURRENT APPLICATION NUMBER: US/10/503,135
; CURRENT FILING DATE: 2004-08-11
; PRIOR APPLICATION NUMBER: PCT/IB03/01161
; PRIOR FILING DATE: 2003-02-13
; PRIOR APPLICATION NUMBER: GB-0203403.1
; PRIOR FILING DATE: 2002-03-13
; NUMBER OF SEQ ID NOS: 181
; SOFTWARE: SeqWin99, version 1.02
; SEQ ID NO 125
; LENGTH: 1036
; TYPE: PRT
; ORGANISM: Chlamydia trachomatis
US-10-503-135-125

Query Match 19.5%; Score 105; DB 5; Length 1036;
Best Local Similarity 28.3%; Pred. No. 0.015;
Matches 26; Conservative 18; Mismatches 48; Indels 0; Gaps 0;
QY 3 ADESTALPDGAGLVLDGTVTALEAEAGMADRIKRELQELKSTGLDVSRRIRVMSV 62
Db 926 ASAEGVASASFAVAVLDQQLPEIMEGIARELVNKINTWRNRKHLVSDRIARLHA 985
QY 63 PAEREDMARTHDLIAGEIATDPEFADLADG 94
Db 986 PVIQEAFAHKEVCEFTLTTSVSVIDYKEG 1017

RESULT 15

US-10-032-585-7198
; Sequence 7198, Application US/10032585
; Publication No. US20030180953A1
; GENERAL INFORMATION:
; APPLICANT: Terry, Roemer D.
; APPLICANT: Bo, Jjiang
; APPLICANT: Charles, Boone
; APPLICANT: Howard, Bussey
; TITLE OF INVENTION: Gene Disruption Methodologies for Drug Target Discovery
; FILE REFERENCE: 10182-005-999
; CURRENT APPLICATION NUMBER: US/10/032,585
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 8000
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7198
; LENGTH: 1088
; TYPE: PRT
; ORGANISM: Candida albicans
US-10-032-585-7198

Query Match 18.2%; Score 98; DB 4; Length 1088;
Best Local Similarity 35.6%; Pred. No. 0.099;
Matches 21; Conservative 14; Mismatches 24; Indels 0; Gaps 0;
QY 17 LVVLDGTVTALEAEAGMADRIKRELQELKSTGLDVSRRIRVMSVPAEREDMARTHDLI 75
Db 985 LIIIDVNLHPELESEBIAELINRIORLRKAGLANTTDVQVQYRVVKTIDLPKIKD 1043
Search completed: April 14, 2006, 18:40:41
Job time : 38.999 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: April 14, 2006, 18:37:38 ; Search time 4.62788 Seconds
(without alignments)
991.565 Million cell updates/sec

Title: US-10-620-246-4

Sequence: 1 MADPESTRALPDGAGLVVL.....ADLADGVAIGDGVRSIEKTI 108

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 217505 seqs, 42489236 residues

Total number of hits satisfying chosen parameters: 217505

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database:

Published Applications: AA New: *
1: /SIDS5/ptodata/2/pubpaa/US08 NEW PUB.pep: *
2: /SIDS5/ptodata/2/pubpaa/US07 NEW PUB.pep: *
3: /SIDS5/ptodata/2/pubpaa/US07 NEW PUB.pep: *
4: /SIDS5/ptodata/2/pubpaa/US07 NEW PUB.pep: *
5: /SIDS5/ptodata/2/pubpaa/US09 NEW PUB.pep: *
6: /SIDS5/ptodata/2/pubpaa/US10 NEW PUB.pep: *
7: /SIDS5/ptodata/2/pubpaa/US11 NEW PUB.pep: *
8: /SIDS5/ptodata/2/pubpaa/US60_NEW PUB.pep: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	134.5	25.0	1149	7	US-11-079-463-5703 Sequence 5703, App
2	112.5	20.9	1080	6	US-10-506-454-899 Sequence 899, App
3	76.5	14.2	660	7	US-11-033-039-385 Sequence 385, App
4	76.5	14.2	661	7	US-11-155-288-13 Sequence 13, App1
5	76.5	14.2	661	7	US-11-119-502-1 Sequence 1, App11
6	74.5	14.2	662	7	US-11-090-439-9 Sequence 9, App11
7	74.5	13.8	485	7	US-11-188-298-6318 Sequence 6318, App
8	72.5	13.5	435	7	US-11-188-298-8860 Sequence 8860, App
9	71.5	13.3	159	7	US-11-055-822-564 Sequence 564, App
10	71.5	13.3	159	7	US-11-055-822-566 Sequence 566, App
11	70.5	13.1	2910	7	US-11-087-100-2 Sequence 2, App11
12	70.5	13.1	2910	7	US-11-087-084-2 Sequence 2, App11
13	70.5	13.1	2910	7	US-11-087-085-2 Sequence 2, App11
14	67	12.5	328	7	US-11-188-298-8241 Sequence 8241, App
15	66.5	12.4	239	7	US-11-188-298-19411 Sequence 19411, App
16	66.5	12.4	506	7	US-11-188-298-1725 Sequence 1725, App
17	66.5	12.4	859	6	US-10-467-657-6084 Sequence 6084, App
18	66	12.3	963	7	US-11-188-298-15124 Sequence 15124, App
19	65	12.1	552	7	US-11-201-916-22 Sequence 22, App1
20	65	12.1	587	7	US-11-096-568A-10866 Sequence 10866, App
21	64.5	12.0	239	6	US-10-518-019A-9 Sequence 9, App11
22	64.5	12.0	239	7	US-11-165-211-37 Sequence 37, App1
23	64.5	12.0	239	7	US-11-165-226-47 Sequence 47, App1
24	64.5	12.0	258	7	US-11-055-822-390 Sequence 390, App
25	64.5	12.0	359	7	US-11-188-298-10148 Sequence 10148, App

26	64.5	12.0	484	6	US-10-506-454-499 Sequence 499, App
27	64.5	12.0	567	6	US-10-330-773-556 Sequence 556, App
28	64.5	12.0	582	6	US-10-330-773-558 Sequence 558, App
29	64	11.9	373	7	US-11-188-298-21978 Sequence 21978, App
30	64	11.9	376	7	US-11-188-298-14205 Sequence 14205, App
31	64	11.9	376	7	US-11-188-298-14377 Sequence 14377, App
32	64	11.9	485	6	US-10-124-598-43 Sequence 43, App1
33	64	11.9	494	6	US-10-506-454-1412 Sequence 1412, App
34	63.5	11.8	199	6	US-10-793-626-2690 Sequence 2690, App
35	63.5	11.8	239	7	US-11-165-211-38 Sequence 38, App1
36	63.5	11.8	239	7	US-11-165-226-48 Sequence 48, App1
37	63.5	11.8	484	7	US-11-188-298-15773 Sequence 15773, App
38	63.5	11.8	506	7	US-11-188-298-4288 Sequence 4288, App
39	63.5	11.8	506	7	US-11-188-298-4385 Sequence 4385, App
40	63.5	11.8	506	7	US-11-188-298-8239 Sequence 8239, App
41	63.5	11.8	506	7	US-11-188-298-11111 Sequence 11111, App
42	63.5	11.8	546	6	US-10-615-668-6 Sequence 6, App11
43	63.5	11.8	637	7	US-11-188-298-6886 Sequence 6886, App
44	63	11.7	266	7	US-11-074-176-166 Sequence 166, App
45	63	11.7	325	7	US-11-063-343-29 Sequence 29, App1

ALIGNMENTS

```
RESULT 1
US-11-079-463-5703
; Sequence 5703, Application US/11079463
; Publication No. US20060073161A1
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO BACTERIOIDES FR.
; FILE REFERENCE: PAT100-03DIV2
; CURRENT APPLICATION NUMBER: US/11/079,463
; CURRENT FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/128,705
; PRIOR FILING DATE: 1999-04-09
; PRIOR APPLICATION NUMBER: US 09/540,209
; PRIOR FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 10444
; SEQ ID NO 5703
; LENGTH: 1149
; TYPE: PRT
; ORGANISM: B. fragilis
US-11-079-463-5703

Query Match      25.0%; Score 134.5; DB 7; Length 1149;
Best Local Similarity 28.7%; Pred. No. 7.6e-06;
Matches 27; Conservative 24; Mismatches 38; Indels 5; Gaps 1;

Cy      13 DGAAGVLDGTYAELAEAGMAKDRIRHLEOKRSGVSDRIRVMSVPAEREDMART 72
Db      1050 EGGTVALAEVTVDELRRRGIABELVNRIOIKRSGSFETIDKIKLTSKPTDAVNE 1109

Cy      73 HRDLAGEIILADTFEFADLADGVAIGDGVRSIE 106
Db      1110 YNSYICNVLTGTSLTAD-----EYVDGTIELNFD 1138

RESULT 2
US-10-506-454-899
; Sequence 899, Application US/10506454
; Publication No. US20060068386A1
; GENERAL INFORMATION:
; APPLICANT: Stebarez, Alexi I
; APPLICANT: Mezhevaya, Katja V
; APPLICANT: Polushin, Nikolai N
; APPLICANT: Shcherbina, Olga V
; APPLICANT: Shakhova, Vera V
; APPLICANT: Malikh, Andrei G
; APPLICANT: Kozavkin, Sergei A
; TITLE OF INVENTION: The Complete Genome and Protein Sequences of the Hyperthermophil.
```

```

; TITLE OF INVENTION: Methanopyrus Kandleri AV19 and Monophyly of Archaeal Methanogens
; TITLE OF INVENTION: and Methods of Use Thereof
; FILE REFERENCE: FID001
; CURRENT APPLICATION NUMBER: US/10/506,454
; CURRENT FILING DATE: 2004-08-31
; PRIOR APPLICATION NUMBER: PCT/US03/06664
; PRIOR FILING DATE: 2003-03-04
; PRIOR APPLICATION NUMBER: 60/361,742
; PRIOR FILING DATE: 2002-03-04
; NUMBER OF SEQ ID NOS: 1722
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 899
; LENGTH: 1080
; TYPE: PRF
; ORGANISM: Methanopyrus kandleri
US-10-506-454-899

Query Match          20.2%; Score 112.5; DB 6; Length 1080;
Best Local Similarity 30.6%; Pred. No. 0.0014;
Matches 26; Conservative 18; Mismatches 34; Indels 7; Gaps 1;

QY 8 TAAAPG-----AGIVVDGTVTALEBAAGAKDRIRRELSTGSLDVSRIYVM 60
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 964 TEDLPGEWEAEPEEGGVVVFVELDEBLSEAWAREYVRKQERKELDLNLEERIKWT 1023
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
QY 61 SVPAEREDWARTRDLIAGEILATD 85
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 1024 ETDEEFAEVEHSEYVRGETRADE 1048
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

RESULT 3
US-11-033-039-385
; Sequence 385, Application US/11033039
; Publication No. US20060002947A1
; GENERAL INFORMATION:
; APPLICANT: HOMPHREYS, ROBERT
; APPLICANT: XU, MINZHEN
; TITLE OF INVENTION: LI-KEY/ANTIGENIC EPI TOPE HYBRID PEPTIDE VACCINES
; FILE REFERENCE: REH-2017US01
; CURRENT APPLICATION NUMBER: US/11/033,039
; CURRENT FILING DATE: 2005-01-11
; PRIOR APPLICATION NUMBER: 10/245,871
; PRIOR FILING DATE: 2002-09-17
; PRIOR APPLICATION NUMBER: 10/197,000
; PRIOR FILING DATE: 2002-07-17
; PRIOR APPLICATION NUMBER: 09/396,813
; PRIOR FILING DATE: 1999-09-14
; NUMBER OF SEQ ID NOS: 1452
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 385
; LENGTH: 660
; TYPE: PRF
; ORGANISM: Homo sapiens
US-11-033-039-385

Query Match          14.2%; Score 76.5; DB 7; Length 660;
Best Local Similarity 25.2%; Pred. No. 4.5;
Matches 30; Conservative 16; Mismatches 44; Indels 29; Gaps 5;

QY 5 PESTALPDGAGIVVDGTVTALEBAAGAKDRIRREL-----QELRKST 48
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 394 PEATGMPAEVSIIVLSGTTAAQVTTTEWETTARELPPEBPGPDASSIMSTESITGSL 453
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
QY 49 G--LDVSDRIYV--MSVPAE-----REDWARTRDLIAGEILATDPEFADLADGVAIGDG 100
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 454 GPLLDGTATRLVKRQVPLDCLYRGSFSVTLDIVG-----TESAEILQAVPSGEG 506
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

RESULT 4
US-11-155-288-13
; Sequence 13, Application US/11155288
; Publication No. US20060008468A1
; GENERAL INFORMATION:
```

```

; APPLICANT: Chhang, Chih-Sheng
; APPLICANT: Simard, John J.L.
; TITLE OF INVENTION: COMBINATIONS OF TUMOR-ASSOCIATED
; TITLE OF INVENTION: ANTIGENS IN DIAGNOSTICS FOR VARIOUS TYPES OF CANCERS
; FILE REFERENCE: MANRK 050A
; CURRENT APPLICATION NUMBER: US/11/155,288
; CURRENT FILING DATE: 2005-06-17
; PRIOR APPLICATION NUMBER: 60/580,969
; PRIOR FILING DATE: 2004-06-17
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: PaetsEQ for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 661
; TYPE: PRF
; ORGANISM: Homo sapiens
US-11-155-288-13

Query Match          14.2%; Score 76.5; DB 7; Length 661;
Best Local Similarity 25.2%; Pred. No. 4.5;
Matches 30; Conservative 16; Mismatches 44; Indels 29; Gaps 5;

QY 5 PESTALPDGAGIVVDGTVTALEBAAGAKDRIRREL-----QELRKST 48
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 395 PEATGMPAEVSIIVLSGTTAAQVTTTEWETTARELPPEBPGPDASSIMSTESITGSL 454
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
QY 49 G--LDVSDRIYV--MSVPAE-----REDWARTRDLIAGEILATDPEFADLADGVAIGDG 100
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 455 GPLLDGTATRLVKRQVPLDCLYRGSFSVTLDIVG-----TESAEILQAVPSGEG 507
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

RESULT 5
US-11-119-502-1
; Sequence 1, Application US/1119502
; Publication No. US20060014668A1
; GENERAL INFORMATION:
; APPLICANT: Genzyme Corporation
; APPLICANT: Nicolette, Charles A.
; TITLE OF INVENTION: Antigenic sp 100 Compounds for Therapy and Diagnosis and Methods
; FILE REFERENCE: 5247PCT
; CURRENT APPLICATION NUMBER: US/11/119,502
; CURRENT FILING DATE: 2005-04-29
; PRIOR APPLICATION NUMBER: 60/422,620
; PRIOR FILING DATE: 2002-10-30
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 661
; TYPE: PRF
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (209)...(217)
US-11-119-502-1

Query Match          14.2%; Score 76.5; DB 7; Length 661;
Best Local Similarity 25.2%; Pred. No. 4.5;
Matches 30; Conservative 16; Mismatches 44; Indels 29; Gaps 5;

QY 5 PESTALPDGAGIVVDGTVTALEBAAGAKDRIRREL-----QELRKST 48
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 395 PEATGMPAEVSIIVLSGTTAAQVTTTEWETTARELPPEBPGPDASSIMSTESITGSL 454
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
QY 49 G--LDVSDRIYV--MSVPAE-----REDWARTRDLIAGEILATDPEFADLADGVAIGDG 100
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 455 GPLLDGTATRLVKRQVPLDCLYRGSFSVTLDIVG-----TESAEILQAVPSGEG 507
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

RESULT 6
US-11-030-439-9
; Sequence 9, Application US/11090439
; Publication No. US2005026442A1
; GENERAL INFORMATION:
```



```
; APPLICANT: Squillace, Rachel
; APPLICANT: Weiner, Michael P.
; TITLE OF INVENTION: Immortalized Human Tubercous Sclerosis Null
; TITLE OF INVENTION: Angiomyolipoma Cell and Method of Use Thereof
; FILE REFERENCE: 24318-502
; CURRENT APPLICATION NUMBER: US/11/090,439
; CURRENT FILING DATE: 2005-03-25
; PRIOR APPLICATION NUMBER: 60/556,344
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 662
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-090-439-9

Query Match          14.2%; Score 76.5; DB 7; Length 662;
Best Local Similarity 25.2%; Pred. No. 4.5;
Matches 30; Conservative 16; Mismatches 44; Indels 29; Gaps 5;

QY 5 PESTALPDGAGLVLDGTVTALEAGMAKDRIRL-----QELAKST 48
DB 395 PRATGMPAEVSIVLSGTTAAQVTTTETWETARLPIPEBGPASSIMSTETIGSL 454
QY 49 G--LDVSDRIRV--MSVPAB-----REDMARTHDLIAGEIILATDFEPADLADGVAIGDG 100
DB 455 GPLDGNATRLVRQVPLDCVLYRGSPSVITDIVG-----IESAEILQAVPSSEG 507

RESULT 7
US-11-188-298-6318
; Sequence 6318, Application US/11188298
; Publication No. US20060075522A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: GENES AND USES FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-21(53452)B
; CURRENT APPLICATION NUMBER: US/11/188,298
; CURRENT FILING DATE: 2005-07-22
; PRIOR APPLICATION NUMBER: 60/592,978
; PRIOR FILING DATE: 2004-07-31
; NUMBER OF SEQ ID NOS: 22569
; SEQ ID NO 6318
; LENGTH: 485
; TYPE: PRT
; ORGANISM: Geobacter metallireducens
US-11-188-298-6318

Query Match          13.8%; Score 74.5; DB 7; Length 485;
Best Local Similarity 28.0%; Pred. No. 4.9;
Matches 33; Conservative 16; Mismatches 56; Indels 13; Gaps 4;

QY 3 ADESTALPDGAGLVLD-----GTVTALEAGMAKDRIRLEQLR-----KSTGLD 51
DB 361 AREVIAALDKINDHYNLRLQAACVALLRDQAVLSECCRRIRRETEWTTLELRSGYD 420
QY 52 V-SDRIRVMSVPAERDMARTHDLIAGEIILATDFEPADLADGVAIGDGVRSIEXT 108
DB 421 VTPSGNYLFAETPPDR-DGKRVYDGLYARKVLVHFDPDLAIGMRISIGTREMECT 477

RESULT 8
US-11-188-298-8860
; Sequence 8860, Application US/11188298
; Publication No. US20060075522A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: GENES AND USES FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-21(53452)B
; CURRENT APPLICATION NUMBER: US/11/188,298
; CURRENT FILING DATE: 2005-07-22
; PRIOR APPLICATION NUMBER: 60/592,978
```

```
; PRIOR FILING DATE: 2004-07-31
; NUMBER OF SEQ ID NOS: 22569
; SEQ ID NO 8860
; LENGTH: 435
; TYPE: PRT
; ORGANISM: Methanothermobacter thermautotrophicus str. Delta H
US-11-188-298-8860

Query Match          13.5%; Score 72.5; DB 7; Length 435;
Best Local Similarity 31.8%; Pred. No. 6.9;
Matches 35; Conservative 12; Mismatches 34; Indels 29; Gaps 8;

QY 2 AADPESTALPDGAGLVLDGTV-----TABLEAGMAKDRIRLEQLRKSTGLDVS 54
DB 85 AEDPE-----LPD-VSISTGTVASIIDYRTGAVHPAFTADLLRANPEL-----LDIAN 134
QY 55 -RIRVMSVPAER---EDMARTHDLIAGEIILATDFEPADLADGVAIGDG 100
DB 135 IRGRAVFNILSENKPEYVETAR-AVYGEI-----KGDAGVVAHG 176

RESULT 9
US-11-055-822-564
; Sequence 564, Application US/11055822
; Publication No. US20050260707A1
; GENERAL INFORMATION:
; APPLICANT: Pompeius, Marcus
; APPLICANT: Kroger, Burkhard
; APPLICANT: Schröder, Hartwig
; APPLICANT: Schöder, Oskar
; APPLICANT: Habertbauer, Gregor
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING
; FILE REFERENCE: BGI-121CPN
; CURRENT APPLICATION NUMBER: US/11/055,822
; CURRENT FILING DATE: 2005-02-11
; PRIOR APPLICATION NUMBER: 09/606,740
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: 60/141,031
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: 60/142,101
; PRIOR FILING DATE: 1999-07-02
; PRIOR APPLICATION NUMBER: 60/148,613
; PRIOR FILING DATE: 1999-08-12
; PRIOR APPLICATION NUMBER: 60/187,970
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: DE 19930476.9
; PRIOR FILING DATE: 1999-07-01
; PRIOR APPLICATION NUMBER: DE 19931415.2
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931418.7
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931419.5
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931420.9
; PRIOR FILING DATE: 1999-07-08
; Remaining Prior Application data removed - See File Wrapper or PAM.
; NUMBER OF SEQ ID NOS: 1158
; SEQ ID NO 564
; LENGTH: 159
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-11-055-822-564

Query Match          13.3%; Score 71.5; DB 7; Length 159;
Best Local Similarity 27.1%; Pred. No. 2.4;
Matches 32; Conservative 20; Mismatches 45; Indels 21; Gaps 6;

QY 1 MAADPESTALPDGAGLVLDGTVTALEAGMAKDRIR-REQLRKSTGLDVS-DRI 58
DB 1 MAKEGILPAVELPAGSLGK--AVVTARWNE--ICRLHGHAVDAGAPAGATVSEKRVIG 56
QY 59 VMSVPAEREDMARTHDLIA-----GILLATDFEPADLADGVA-----IGDGV 101
```

DB 57 ALELPVVGELATHTDAVVALGCVVRGCTPHFVDCSVTEGLTRIALDTSTPIGNV 114

RESULT 10

US-11-055-822-566

Sequence 566, Application US/11055822

Publication No. US20050260707A1

GENERAL INFORMATION:

APPLICANT: Pompejus, Markus

APPLICANT: Krogger, Burhard

APPLICANT: Schroder, Hartwig

APPLICANT: Zeider, Oskar

APPLICANT: Haberer, Gregor

TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING

TITLE OF INVENTION: METABOLIC PATHWAY PROTEINS

FILE REFERENCE: BGI-121CPCN

CURRENT APPLICATION NUMBER: US/11/055,822

CURRENT FILING DATE: 2005-02-11

PRIOR APPLICATION NUMBER: 09/606,740

PRIOR FILING DATE: 2000-06-23

PRIOR APPLICATION NUMBER: 60/141,031

PRIOR FILING DATE: 1999-06-25

PRIOR APPLICATION NUMBER: 60/142,101

PRIOR FILING DATE: 1999-07-02

PRIOR APPLICATION NUMBER: 60/148,613

PRIOR FILING DATE: 1999-08-12

PRIOR APPLICATION NUMBER: 60/187,970

PRIOR FILING DATE: 2000-03-09

PRIOR APPLICATION NUMBER: DE 19930476.9

PRIOR FILING DATE: 1999-07-01

PRIOR APPLICATION NUMBER: DE 19931415.2

PRIOR FILING DATE: 1999-07-08

PRIOR APPLICATION NUMBER: DE 19931418.7

PRIOR FILING DATE: 1999-07-08

PRIOR APPLICATION NUMBER: DE 19931419.5

PRIOR FILING DATE: 1999-07-08

PRIOR APPLICATION NUMBER: DE 19931420.9

PRIOR FILING DATE: 1999-07-08

Remaining prior Application data removed - See file Wrapper or PALM.

NUMBER OF SEQ ID NOS: 1158

SEQ ID NO 566

LENGTH: 159

TYPE: PRT

ORGANISM: Corynebacterium glutamicum

US-11-055-822-566

Query Match 13.3%; Score 71.5; DB 7; Length 159;

Best Local Similarity 27.1%; Pred. No. 2.4;

Matches 32; Conservative 20; Mismatches 45; Indels 21; Gaps 6;

DB 1 MAADPESTALPDGAGLVVDGVTVALEAEGWAKORI-RELOELRKSTGLDVS-D-RIRV 58

1 MAEGEPAYVELPDAAGSLKV-AVVTAKMNE--ICRLHGAADAGAAATVSEYVIG 56

DB 59 VMSVPAEREDWARTHDLA-----GELATDFEFADLADGVA-----IGDGV 101

57 ALELPVVGELATHTDAVVALGCVVRGCTPHFVDCSVTEGLTRIALDTSTPIGNV 114

RESULT 11

US-11-087-100-2

Sequence 2, Application US/11087100

Publication No. US2005026440A1

GENERAL INFORMATION:

APPLICANT: Metz, James

APPLICANT: Barclay, William

APPLICANT: Platt, James

APPLICANT: Kuner, Jerry

TITLE OF INVENTION: Nucleic Acid Molecule Encoding ORFA of a PUFA Polyketide Synthase

TITLE OF INVENTION: System and Uses Thereof

FILE REFERENCE: 2997-29

CURRENT APPLICATION NUMBER: US/11/087,100

CURRENT FILING DATE: 2005-03-21

PRIOR APPLICATION NUMBER: 09/231,899

PRIOR FILING DATE: 1999-01-14

PRIOR APPLICATION NUMBER: 60/284,066

PRIOR FILING DATE: 2001-04-16

PRIOR APPLICATION NUMBER: 60/298,796

PRIOR FILING DATE: 2001-06-15

PRIOR APPLICATION NUMBER: 60/323,269

PRIOR FILING DATE: 2001-09-18

NUMBER OF SEQ ID NOS: 37

SOFTWARE: PatentIn version 3.3

SEQ ID NO 2

LENGTH: 2910

TYPE: PRT

ORGANISM: Schizochytrium sp.

US-11-087-084-2

Query Match 13.1%; Score 70.5; DB 7; Length 2910;

Best Local Similarity 30.4%; Pred. No. 1.3e+02;

Matches 31; Conservative 13; Mismatches 49; Indels 9; Gaps 4;

DB 2 AADPESTALPDGAGLVVDGVTVALEA--EGWAKDRIRELOELRKSTGLDVS-D-RIRV 58

1877 AANPAPAAAPAVSSLEAEVVMVEVLAAGTGYETDMIESDMELETETGIDISIKRVEI 1936

DB 59 VMSVPA-----EREDWARTHDLIAGEILATDFEFADLADGVA 96

1937 LSEVQAMLNEAKVDVALSRTTVGSEV--DAWKAKIAGGSA 1976

RESULT 12

US-11-087-084-2

Sequence 2, Application US/11087084

Publication No. US20050273883A1

GENERAL INFORMATION:

APPLICANT: Metz, James

APPLICANT: Barclay, William

APPLICANT: Platt, James

APPLICANT: Kuner, Jerry

TITLE OF INVENTION: Nucleic Acid Molecule Encoding ORFA of a PUFA Polyketide Synthase

TITLE OF INVENTION: System and Uses Thereof

FILE REFERENCE: 2997-29

CURRENT APPLICATION NUMBER: US/11/087,084

CURRENT FILING DATE: 2005-03-21

PRIOR APPLICATION NUMBER: 09/231,899

PRIOR FILING DATE: 1999-01-14

PRIOR APPLICATION NUMBER: 60/284,066

PRIOR FILING DATE: 2001-04-16

PRIOR APPLICATION NUMBER: 60/298,796

PRIOR FILING DATE: 2001-06-15

PRIOR APPLICATION NUMBER: 60/323,269

PRIOR FILING DATE: 2001-09-18

NUMBER OF SEQ ID NOS: 37

SOFTWARE: PatentIn version 3.3

SEQ ID NO 2

LENGTH: 2910

TYPE: PRT

ORGANISM: Schizochytrium sp.

US-11-087-084-2

Query Match 13.1%; Score 70.5; DB 7; Length 2910;

Best Local Similarity 30.4%; Pred. No. 1.3e+02;

Matches 31; Conservative 13; Mismatches 49; Indels 9; Gaps 4;

DB 2 AADPESTALPDGAGLVVDGVTVALEA--EGWAKDRIRELOELRKSTGLDVS-D-RIRV 58

1877 AANPAPAAAPAVSSLEAEVVMVEVLAAGTGYETDMIESDMELETETGIDISIKRVEI 1936

DB 59 VMSVPA-----EREDWARTHDLIAGEILATDFEFADLADGVA 96

1937 LSEVQAMLNEAKVDVALSRTTVGSEV--DAWKAKIAGGSA 1976

RESULT 13
US-11-087-085-2
; Sequence 2, Application US/11087085
; Publication No. US20050273884A1
; GENERAL INFORMATION:
; APPLICANT: Metz, James
; APPLICANT: Barclay, William
; APPLICANT: Platt, James
; APPLICANT: Kumer, Jerry
; TITLE OF INVENTION: Nucleic Acid Molecule Encoding ORF of a Pura polypeptide Synthase
; TITLE OF INVENTION: System and Uses Thereof
; FILE REFERENCE: 2997-29
; CURRENT APPLICATION NUMBER: US/11/087,085
; CURRENT FILING DATE: 2005-03-21
; PRIOR FILING DATE: 09/231,899
; PRIOR FILING DATE: 1999-01-14
; PRIOR APPLICATION NUMBER: 60/284,066
; PRIOR FILING DATE: 2001-04-16
; PRIOR APPLICATION NUMBER: 60/298,796
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/323,269
; PRIOR FILING DATE: 2001-09-18
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 2
; LENGTH: 2910
; TYPE: PRT
; ORGANISM: Schizochytrium sp.
US-11-087-085-2

Query Match 13.1%; Score 70.5; DB 7; Length 2910;
Best Local Similarity 30.4%; Pred. No. 1.3e+02;
Matches 31; Conservative 13; Mismatches 49; Indels 9; Gaps 4;
QY 2 AADPESTALPD-GAGLVLDGVTVALEA-BGMADRIRELOELKSTGLDVSRIIV 58
DB 1877 AAPAPAPAAAPAVASSELKRAETVMEVLAKTGYETDMIESDMELETIGIDISIKRVEI 1936
QY 59 VMSVPA-----EREDMARTHRDLIAGRIATDPFPAIDAGVA 96
DB 1937 LSEVQAMLVKAVDVALSTRIVGEV--DAKKAELIAGGSA 1976

RESULT 14
US-11-188-298-8241
; Sequence 8241, Application US/11188298
; Publication No. US20060075522A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: GENES AND USES FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-21(53452)B
; CURRENT APPLICATION NUMBER: US/11/188,298
; CURRENT FILING DATE: 2005-07-22
; PRIOR APPLICATION NUMBER: 60/592,978
; PRIOR FILING DATE: 2004-07-31
; NUMBER OF SEQ ID NOS: 22569
; SEQ ID NO 8241
; LENGTH: 328
; TYPE: PRT
; ORGANISM: GIBBERELLA ZEA PH-1
US-11-188-298-8241

Query Match 12.5%; Score 67; DB 7; Length 328;
Best Local Similarity 25.2%; Pred. No. 18;
Matches 26; Conservative 17; Mismatches 42; Indels 18; Gaps 5;
QY 7 STPALPDGAGLVLDGVTVALEABGMAKDRIRLEQL-----RKSTGLDVS 54
DB 28 TTGVSPGCGALFVBAIAVAPELAIAGRNINKQOTADHLASKHPNLTLLTIDLS 87
QY 55 RIRVMSVPAEREDMARTHR--DLIAGE--LLATDFPADIADG 94
DB 88 -LASIRSAAQVVGWSDVHKIDVLVNNAGIMATDFKLT--DG 127

RESULT 15
US-11-188-298-19411
; Sequence 19411, Application US/11188298
; Publication No. US20060075522A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: GENES AND USES FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-21(53452)B
; CURRENT APPLICATION NUMBER: US/11/188,298
; CURRENT FILING DATE: 2005-07-22
; PRIOR APPLICATION NUMBER: 60/592,978
; PRIOR FILING DATE: 2004-07-31
; NUMBER OF SEQ ID NOS: 22569
; SEQ ID NO 19411
; LENGTH: 239
; TYPE: PRT
; ORGANISM: Magnetospirillum magnetotacticum
US-11-188-298-19411

Query Match 12.4%; Score 66.5; DB 7; Length 239;
Best Local Similarity 26.8%; Pred. No. 14;
Matches 37; Conservative 14; Mismatches 50; Indels 37; Gaps 5;
QY 5 PSTPALPDGAGLVLDGVTVALEABGMAKDRIRLEOKSTGLD-----VSDRI-RV 58
DB 29 PELAGVDPAFGVLVIDGDSRA---AVGGADIPFISGISISKVTLLALGWTGDRIMR 85
QY 59 VMSVPA-----EREDMARTHRDLIAGRIATDPFPAIDAGVA 96
DB 86 VGREPSPSPNSIVQLERENGIRPNPFINAGAIATVDVILSGHPREALGRIIFMQFLA 145
QY 91 LADGVAIGCVRSIKET 108
DB 146 QDNGIAIDKVAASERT 163

Search completed: April 14, 2006, 18:42:18
Job time : 5.62788 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using SW model

Run on: April 14, 2006, 17:03:40 ; Search time 63.931 Seconds
(without alignments)
1133.996 Million cell updates/sec

Title: US-10-620-246-8

Perfect score: 831
Sequence: 1 MAQITLRGNAINTVGBLPAY.....LVPEIAQBPVYKALALGA 165

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_21.*
1: geneseqp1980s.*
2: geneseqp1990s.*
3: geneseqp2000s.*
4: geneseqp2001s.*
5: geneseqp2002s.*
6: geneseqp2003as.*
7: geneseqp2003bs.*
8: geneseqp2004s.*
9: geneseqp2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	831	100.0	165	2	AAW72887 Mycobacte
2	831	100.0	165	2	AAW72887 Mycobacte
3	831	100.0	165	5	AAW50737 Mycobacte
4	514.5	61.9	170	4	AAU61969 Propionib
5	514.5	61.9	170	6	ABM58488 Pseudomon
6	486	58.5	194	7	ABO82493 Pseudomon
7	473.5	57.0	165	4	AAAG90953 C. glutami
8	466.5	56.1	165	4	AAAB78902 C. glutami
9	466.5	56.1	165	4	AAAB78902 C. glutami
10	441	53.1	172	7	ADP04323 Bacteri
11	418.5	50.4	176	7	ABO61199 Klebsiell
12	417.5	50.2	168	4	AAU29352 Novel mar
13	406	48.9	168	6	ABM69123 Phototrab
14	341.5	41.1	123	4	AAAB78109 Coryneb
15	340.5	41.0	165	5	ABBA8059 Listeria
16	320.5	38.6	167	6	ABU02121 S. pneumo
17	320.5	38.6	194	2	AAV11220 S. pneumo
18	319.5	38.4	172	8	ADK46702 Streptoco
19	319.5	38.4	201	8	ADR96491 Novel S.
20	319.5	38.4	201	6	AAAG6361 Streptoco
21	314	37.8	170	6	ABM75338 Streptoco
22	314	37.8	173	2	AAW89770 Staphyloc
23	314	37.8	175	9	AAEA22101 Campyloba
24	307	36.9	164	4	AAAG81579 S. epider

25	307	36.9	168	5	ABP38919	Abp38919 Staphyloc
26	307	36.9	168	8	ADSO6624	Adso6624 Staphyloc
27	302	36.3	171	9	ABM96084	Abm96084 M. xanthu
28	282.5	34.0	144	2	AAW55084	Aaw55084 Streptoco
29	282.5	34.0	144	5	ABP54578	Abp54578 S. pneumo
30	282.5	34.0	144	7	ADCA5125	Adca5125 S. pneumo
31	281.5	33.9	164	8	ADVB8857	Advb8857 Streptoco
32	281.5	33.9	164	8	ADVB80110	Advb80110 Streptoco
33	281.5	33.9	164	8	ADVB82243	Advb82243 Streptoco
34	275.5	33.2	164	5	ABP28104	Abp28104 Streptoco
35	228.5	27.5	160	5	ABBS5610	Abbs5610 Lactococc
36	225	27.1	126	8	ADM91680	Adm91680 Lactococc
37	195.5	23.5	108	2	AAW77683	Aaw77683 Tagd prot
38	181.5	21.8	110	7	ADH87845	Adh87845 Enterococ
39	154.5	18.6	274	9	ABM97519	Abm97519 M. xanthu
40	138.5	16.7	164	8	ADSA4517	Adsa4517 Bacteri
41	122.5	14.7	257	5	ABP53043	Abp53043 Rat AOP-1
42	122.5	14.7	257	7	ADDA7422	Adda7422 Rat Prote
43	121.5	14.6	257	5	ADBS8527	Adbs8527 Rat Prote
44	121.5	14.6	257	7	ABP53044	Abp53044 Mouse AOP
45	121.5	14.6	257	7	ABR61585	AbR61585 Murine pe

ALIGNMENTS

RESULT 1	
AAW72887	
ID	AAW72887 standard; protein; 165 AA.
AC	AAW72887;
DT	21-JAN-1999 (first entry)
XX	
DB	Mycobacterium tuberculosis antigen CFP20.
XX	
KW	Mycobacterium tuberculosis; antigen; vaccine; immunological; immunogen; infection.
XX	
OS	Mycobacterium tuberculosis.
XX	
PN	W09844119-A1.
XX	
PD	08-OCT-1998.
XX	
PF	01-APR-1998; 98WO-DK000132.
XX	
PR	02-APR-1997; 97DK-00000376.
XX	
PR	18-APR-1997; 97US-0044624P.
XX	
PR	10-NOV-1997; 97DK-00001277.
XX	
PR	05-JAN-1998; 98US-0070488P.
XX	
PA	(STAT-) STATENS SERUM INST.
XX	
PI	Andersen P, Nielsen R, Rosenkrands I, Weldingh K, Rasmussen PB, Oettinger T, Florio W,
XX	
DR	N-PSDB; AAV63918.
XX	
PT	New isolated mycobacteria polypeptides and nucleic acids - used for developing products for the diagnosis of or vaccination against mycobacterial infections, particularly tuberculosis.
XX	
PS	Claim 1; Page 129-130; 163pp; English.
XX	
CC	The present sequence represents a Mycobacterium tuberculosis protein. Products from the present invention, which describes protein fragments and nucleic acid fragments derived from M. tuberculosis, can be used in the detection of and prevention of mycobacterial infections. In particular, the proteins and nucleic acids can be used for the diagnosis of or vaccination against tuberculosis caused by M. tuberculosis, M. africanum or M. bovis
CC	

XX Sequence 165 AA;
 SQ
 Query Match 100.0%; Score 831; DB 2; Length 165;
 Best Local Similarity 100.0%; Pred. No. 1.1e-83;
 Matches 165; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 1 MAQITLRGNAINTVGLPAVGSPPAPFTLTGDLGVISDQFRGKSVLTNIPPSVDTPTVC 60
 1 MAQITLRGNAINTVGLPAVGSPPAPFTLTGDLGVISDQFRGKSVLTNIPPSVDTPTVC 60
 61 ATSVRTFDBRAAASGATVLCVSKDLPFAQKRCGAGGTENWPASAFRDSFGEDYGVTTA 120
 61 ATSVRTFDBRAAASGATVLCVSKDLPFAQKRCGAGGTENWPASAFRDSFGEDYGVTTA 120
 121 DCPMAGLLARAIIVIGADGNVAYTELVPEIAQEPNTEAALALGA 165
 121 DCPMAGLLARAIIVIGADGNVAYTELVPEIAQEPNTEAALALGA 165

RESULT 2

AAV21904
 ID AAV21904 standard; protein, 165 AA.

AAV21904;
 AC

06-SEP-1999 (first entry)
 DT

Amino acid sequence of antigen CFP20.
 XX

Immunogenic; Mycobacterium tuberculosis; immune response; infection;
 tuberculosis; fusion polypeptide; T-cell epitope; BSAT-6; MPT59; TB;
 pharmaceutical; vaccination; M. africanum; M. bovis; CFP7A; CFP30A;
 CFP7B; CFP19; CFP27; CFP30A; RDI-ORF; CFP10A; CFP16; CFP19; CFP23;
 CFP25A; CFP30B; CFP7B.

XX Mycobacterium tuberculosis.
 OS

MO9924577-A1.
 XX

20-MAY-1999.
 XX

08-OCT-1998; 98WO-DK000438.
 XX

10-NOV-1997; 97DK-00001277.
 XX

05-JAN-1998; 98US-0070488P.
 XX

01-APR-1998; 98WO-DK000132.
 XX

(STAT-) STATENS SERUM INST.
 PA

Anderesen P, Skjot R;
 PI

WPI, 1999-347282/29.
 XX

N-PSDB; AAX81004.
 DR

New immunogenic fragment of Mycobacterium tuberculosis.
 PT

Example 3, Page 60; 265pp; English.
 XX

The invention describes a substantially pure immunogenic polypeptide
 fragment (1) from Mycobacterium tuberculosis that is able to evoke a
 protective immune response against infections by mycobacteria belonging
 to the tuberculosis complex. The invention provides a (1) fusion
 polypeptide comprising at least one polypeptide fragment (1) and at least
 one fusion partner; (2) a fusion polypeptide fragment comprising a T-cell
 epitope from M. tuberculosis protein BSAT-6, or MPT59 and a second
 different amino acid sequence from M. tuberculosis, and/or including a
 sequence which protects the first amino acid sequence from in vivo
 degradation or post-translational processing; (3) a nucleic acid fragment
 that encodes the above polypeptides. The polypeptides and nucleic acid
 are useful as pharmaceuticals, for diagnosis of and as antigens for
 vaccination against TB caused by Mycobacterium tuberculosis, africanum or
 bovis. The polypeptides are also useful for diagnosing ongoing or

CC previous sensitization in an animal with bacteria belonging to the
 CC tuberculosis complex. The invention also describes the use of CFP7A or
 CC CFP30A or a T-cell epitope of for the induction of a strong immune
 CC response in a mammal; use of CFP7B, CFP19 or MPT59-BSAT6 or a T-cell
 CC epitope of for diagnosis of TB in a mammal by performing a DTH type skin
 CC test; use of CFP27, CFP30A, RDI-ORF2, RDI-ORF3, RDI-ORF5, MPT59-BSAT6,
 CC BSAT6-MPT59, CFP10A, CFP16, CFP19, CFP23, CFP25A, CFP30B, CFP7B or a T-
 CC cell epitope of for the preparation of an immunological composition; and
 CC for the preparation of a subunit vaccine
 XX

SQ Sequence 165 AA;

Query Match 100.0%; Score 831; DB 2; Length 165;
 Best Local Similarity 100.0%; Pred. No. 1.1e-83;
 Matches 165; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MAQITLRGNAINTVGLPAVGSPPAPFTLTGDLGVISDQFRGKSVLTNIPPSVDTPTVC 60
 1 MAQITLRGNAINTVGLPAVGSPPAPFTLTGDLGVISDQFRGKSVLTNIPPSVDTPTVC 60
 61 ATSVRTFDBRAAASGATVLCVSKDLPFAQKRCGAGGTENWPASAFRDSFGEDYGVTTA 120
 61 ATSVRTFDBRAAASGATVLCVSKDLPFAQKRCGAGGTENWPASAFRDSFGEDYGVTTA 120

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

RESULT 3

AAV50737
 ID AAV50737 standard; protein, 165 AA.

AAV50737;
 AC

18-APR-2002 (first entry)
 DT

Mycobacterium tuberculosis immunodominant Mtb protein Tpx.
 XX

Mtb; Tpx; immunogen; mycobacteria; immunisation; vaccine.
 XX

Mycobacterium tuberculosis.
 OS

WO200204018-A2.
 XX

17-JAN-2002.
 XX

10-JUL-2001; 2001WO-US021717.
 XX

10-JUL-2000; 2000US-0217646P.
 XX

(COLS) UNIV COLORADO STATE RES FOUND.
 PA

Orme IM, Bellisle JT;
 PI

WPI; 2002-164602/21.
 XX

Vaccine for boosting immunity to mycobacteria when administered in mid-
 PT life in a subject who has been vaccinated in childhood with Bacillus
 PT Calmette-Guerin, has purified proteins from mycobacterium tuberculosis.
 XX

Claim 8; Page 18; 61pp; English.
 XX

The present sequence is that of the Mycobacterium tuberculosis (Mtb)
 strain H37Rv gene Rv1932 product, designated Tpx. This is one of 31
 immunodominant secreted or cytosolic Mtb proteins of strain H37Rv (see
 AAM50725-59) discovered through the use of 2-dimensional liquid phase
 electrophoresis coupled with an in vitro interferon-gamma assay and
 liquid chromatography-mass spectrometry. The immunogens stimulate a
 strong interferon-gamma response from T cells of M. tuberculosis infected
 mice. The invention provides vaccine compositions for boosting immunity
 to mycobacteria when administered in mid-life to a subject who has been
 vaccinated neonatally or in early childhood with BCG and in whom

CC protective immunity has waned. The vaccine compositions comprise 1 or
 CC more of the 31 purified immunogenic proteins. When used as immunogens,
 CC the secreted Mtb proteins lack the secreted signal sequence. A preferred
 CC protein is Ag85A (see AAM50759), the secreted product of the Rv3084v gene
 XX

SO Sequence 165 AA;

Query Match 100.0%; Score 831; DB 5; Length 165;
 Best Local Similarity 100.0%; Pred. No. 1.1e-83;
 Matches 165; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAQITLRGNAINVTGELPAVGSPPAPFTLTGDLGVISDQFRGKSVTLNIFPSVDTPVC 60
 DB 1 MAQITLRGNAINVTGELPAVGSPPAPFTLTGDLGVISDQFRGKSVTLNIFPSVDTPVC 60

QY 61 ATSVRTFDEPRAAASGATVLCVSKDLPPAQRFCGAGETENMPASAFRDSFGEDYVTI 120
 DB 61 ATSVRTFDEPRAAASGATVLCVSKDLPPAQRFCGAGETENMPASAFRDSFGEDYVTI 120

QY 121 DGPMAGLARAIIVTIGADGNVAYTELVPETIAQEPNVEALALGA 165
 DB 121 DGPMAGLARAIIVTIGADGNVAYTELVPETIAQEPNVEALALGA 165

RESULT 4
 AAU61969 standard; protein; 170 AA.

AC AAU61969;
 DT 27-FEB-2002 (first entry)

DE Propionibacterium acnes immunogenic protein #22865.

KM SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
 KM uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
 KM inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
 KM dermatological; osteopathic; neuroprotectant.

OS Propionibacterium acnes.

PN WO200181581-A2.

PD 01-NOV-2001.

PF 20-APR-2001; 2001WO-US012865.

PR 21-APR-2000; 2000US-019047P.

PR 02-JUN-2000; 2000US-020841P.

PR 07-JUL-2000; 2000US-0216747P.

PA (CORI-) CORIXA CORP.

PI Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;

PI L'maisonneuve J, Zhang Y, Jen S, Carter D;

DR WPI; 2001-616774/71.

DR N-PSDB; AAS59621.

PT Propionibacterium acnes polypeptides and nucleic acids useful for
 PT vaccinating against and diagnosing infections, especially useful for
 PT treating acne vulgaris.

PS Example 1; SEQ ID NO 23164; 1069pp; English.

CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
 CC polypeptides. The proteins and their associated DNA sequences are used in
 CC the treatment, prevention and diagnosis of medical conditions caused by
 CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
 CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
 CC P. acnes is also involved in infections of bone, joints and the central
 CC nervous system, however it is particularly involved in the inflammatory
 CC lesions associated with acne vulgaris. A method for detecting the

CC presence or absence of P. acnes in a patient comprises contacting a
 CC sample with a binding agent that binds to the proteins of the invention
 CC and determining the amount of bound protein in the sample. The
 CC polypeptides may be used as antigens in the production of antibodies
 CC specific for P. acnes proteins. These antibodies can be used to
 CC downregulate expression and activity of P. acnes polypeptides and
 CC therefore treat P. acnes infections. The antibodies may also be used as
 CC diagnostic agents for determining P. acnes presence, for example, by
 CC enzyme linked immunosorbent assay (ELISA). Note: The sequence data for
 CC this patent did not form part of the printed specification, but was
 CC obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 170 AA;

Query Match 61.9%; Score 514.5; DB 4; Length 170;
 Best Local Similarity 63.2%; Pred. No. 1.5e-48;
 Matches 103; Conservative 21; Mismatches 38; Indels 1; Gaps 1;

QY 1 MAQITLRGNAINVTGELPAVGSPPAPFTLTGDLGVISDQFRGKSVTLNIFPSVDTPVC 60
 DB 4 MATTALMGKRLNVTGDLPGVGSLLPSFALVKSLSRLSDRLKGLVNLNIFPSVDTPVC 63

QY 61 ATSVRTFDEPRAAASGAT-VLCVSKDLPPAQRFCGAGETENMPASAFRDSFGEDYVTI 119
 DB 64 ATSVRTFENKAGLDITIVLCVSRDLPPAQRFCGAGETENMPASAFRSHFGKDLGVTI 123

QY 120 ADGPMAGLARAIIVTIGADGNVAYTELVPETIAQEPNVEALAA 162
 DB 124 ADGPMAHLARAIIVNAGSKVTYQLVDEITTEPDYDALAA 166

RESULT 5
 ABM58488

ID ABM58488 standard; protein; 170 AA.

AC ABM58488;

DT 20-OCT-2003 (first entry)

DE Propionibacterium acnes predicted ORF-encoded polypeptide #23164.

KM Acne vulgaris; antiseborrheic; dermatological; antibacterial;

KM immunostimulant; immune response; vaccine.

OS Propionibacterium acnes.

PN WO2003033515-A1.

PD 24-APR-2003.

PF 11-OCT-2002; 2002WO-US032727.

PR 15-OCT-2001; 2001US-00978825.

PA (CORI-) CORIXA CORP.

PI Mitcham JL, Skeiky YAW, Persing DH, Bhatia A, Maisonneuve JL;

PI Zhang Y, Wang S, Jen S, Lodes MO, Benson DR, Jones R, Carter D;

PI Barth B, Valiave-Douglas J;

DR WPI; 2003-381789/36.

DR N-PSDB; ACF64550.

PT New Propionibacterium acnes polypeptides and polynucleotides encoding the
 PT polypeptide, useful for diagnosing, preventing or treating acne vulgaris,
 PT or for stimulating an immune response specific for a P. acnes protein.

PS Example 1; SEQ ID NO 23164; 1481pp; English.

CC The invention relates to an isolated polynucleotide (ACF64435-ACF64733)
 CC encoding a Propionibacterium acnes protein. The invention also relates to
 CC polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to

CC immunogenic fragments of *P. acnes* polypeptides. The invention
CC additionally encompasses expression vectors and host cells comprising a
CC polynucleotide of the invention; antibodies against polypeptides of the
CC invention; fusion proteins comprising a polypeptide of the invention; a
CC method for stimulating an immune response specific for a *P. acnes*
CC polypeptide and an isolated T cell population comprising T cells prepared
CC via this method; a vaccine composition (comprising *P. acnes* polypeptides,
CC polynucleotides, antibodies, fusion proteins, T cell populations, or
CC antigen-presenting cells that express the polypeptide); a method and kit
CC for detecting or determining the presence or absence of *P. acnes* in a
CC patient; and a method for inhibiting the development of *P. acnes* in a
CC patient. The *P. acnes* polypeptides, polynucleotides, antibodies, fusion
CC proteins, T cell populations or antigen-presenting cells that express the
CC polypeptides are useful for diagnosing, preventing or treating *acne*
CC vulgaris, or for stimulating an immune response specific for a *P. acnes*
CC protein. The polynucleotides can also be used as probes or primers for
CC nucleic acid hybridisation. The vaccine composition is useful for the
CC stimulation of an immune response against *P. acnes*, or for treating *acne*,
CC and the kit is useful for performing a diagnostic assay. The present
CC sequence represents a polypeptide predicted to be encoded by an ORF (open
CC reading frame) contained within the *P. acnes* polynucleotides of the
CC invention. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pat_sequences
CC
CC
SQ Sequence 170 AA;

Query Match 61.9%; Score 514.5; DB 6; Length 170;
Best Local Similarity 63.2%; Pred. No. 1.5e-48; Indels 1; Gaps 1;
Matches 103; Conservative 21; Mismatches 38; Indels 1; Gaps 1;
DB 1 MAQITIRGNAINTVGEIPAVGSPAPAFITLGGDLGVISDQFRGKSVLTNIFPSVDTPTVC 60
4 MATTALMGKRLNVVGDLPQVGSLLPSFALVKSDELRSDELKGLVLTIFPSVDPTGVC 63
QY 61 ATSVRTFDEPRAASGAT-VTCVSKDLFPFAQKRCGAEETENWPAFAFRDSFGEDYGT 119
DB 64 ATSVRTFNEKRAQGLDITIVLCVSRDLFPFAQARFCGAEIGNVVAFAFRSHFGKDLGVT 123
QY 120 ADGPMAGLLARATVIGADGNVATLVEPIAQBPNYEALAA 162
DB 124 ADGPMQHLARATLVVAGSKVITVQLVDEITTEPTDAALEA 166

RESULT 6
ABO82493
ID ABO82493 standard; protein; 194 AA.
AC ABO82493;
XX
XX
DT 29-JUL-2004 (first entry)
XX
DE Pseudomonas aeruginosa polypeptide #14668.
XX
XX Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.
XX
OS Pseudomonas aeruginosa.
XX
XX US551795-B1.
XX
XX 22-APR-2003.
XX
XX 18-FEB-1999; 99US-00252991.
XX
XX 18-FEB-1998; 98US-0074788P.
XX
XX 27-JUL-1998; 98US-0094190P.
XX
XX (GENO-) GENOME THERAPEUTICS CORP.
XX
XX Rubenfield MJ, Nolling J, Deloughery C, Bush D;
XX WPI, 2003-615309/58.
XX
XX N-PSDB; ABD16064.
DR

XX
XX Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
XX useful as molecular targets for diagnostics, prophylaxis and treatment of
XX pathological conditions resulting from bacterial infection.
XX
XX
XX disclosure; SEQ ID NO 31239; 455pp; English.
XX
XX The invention relates to Pseudomonas aeruginosa polypeptides and the
XX polynucleotides encoding them. The sequences are useful in diagnosis and
XX therapy of pathological conditions, as molecular targets for diagnostics,
XX prophylaxis and treatment of pathological conditions resulting from a
XX bacterial infection, for evaluating a compound, such as a polypeptide,
XX for the ability to bind a *P. aeruginosa* nucleic acid, as components of
XX effective antibacterial targets, as targets for antibacterial drugs,
XX including anti-*P. aeruginosa* drugs, as templates for recombinant
XX production of *P. aeruginosa*-derived peptides or polypeptides, as target
XX components for diagnosis and/or treatment of *P. aeruginosa*-caused
XX infection, and in detection of *P. aeruginosa* sequences or other sequences
XX of Pseudomonas species using biochip technology. Sequences ABO67826-
XX ABO84396 represent *P. aeruginosa* polypeptides of the invention. Note: The
XX sequence data for this patent did not form part of the printed
XX specification but was obtained in electronic format from USPTO at
XX seqdata.uspto.gov/sequence.html
XX
XX
SQ Sequence 194 AA;

Query Match 58.5%; Score 486; DB 7; Length 194;
Best Local Similarity 60.0%; Pred. No. 2.7e-45; Indels 2; Gaps 2;
Matches 99; Conservative 20; Mismatches 44; Indels 2; Gaps 2;
QY 1 MAQITIRGNAINTVGEIPAVGSPAPAFITLGGDLGVISDQFRGKSVLTNIFPSVDTPTVC 60
DB 30 MAQITLGNPNVNVGDLPOKGAQAPAFSLVGGDLADVTLENFAGKRVLTNIFPSVDPTPTC 89
QY 61 ATSVRTFDEPRAASGATVTCVSKDLFPFAQKRCGAEETENWPAFAFRDSFGEDYGT 118
DB 90 ATSVRTFNEKRAQGLDITIVLCVSRDLFPFAQKRCGAEIGNVVAFAFRSHFGKDLGVT 149
QY 119 IADGPMAGLLARATVIGADGNVATLVEPIAQBPNYEALAA 163
DB 150 IASGPMAGLLARATLVVAGSKVITVQLVDEITTEPTDAALEA 194

RESULT 7
AAG90953
ID AAG90953 standard; protein; 165 AA.
XX
XX AAG90953;
XX
XX
DT 26-SEP-2001 (first entry)
XX
XX
DE C glutamylum protein fragment SEQ ID NO: 4707.
XX
XX
XX Coryneform bacterium; amino acid synthesis; vitamin; saccharide;
XX organic acid synthesis.
XX
XX Corynebacterium glutamicum.
XX
XX
XX EP1108790-A2.
XX
XX 20-JUN-2001.
XX
XX 18-DEC-2000; 2000EP-00127688.
XX
XX
XX 16-DEC-1999; 99JP-00377484.
XX
XX 07-APR-2000; 2000JP-00159162.
XX
XX 03-AUG-2000; 2000JP-00280988.
XX
XX (KYOW) KYOWA HAKKO KOGYO KK.
XX
XX Nakagawa S, Mizoguchi H, Ando S, Hayaishi M, Ochiai K, Yokoi H;
XX Tateishi N, Senoh A, Ikeda M, Ozaki A;
XX

DR WPI; 2001-376931/40.
XX N-PSDB; AAB66172.
PT Novel polynucleotides derived from Corynebacterium bacteria, for identifying
PT mutation point of a gene, measuring expression of a gene, analyzing
PT expression profile or pattern of a gene and identifying homologous gene.
XX
XX Claim 17; SEQ ID NO 4707; 246bp + Sequence Listing; English.
XX
CC The present invention provides a number of nucleotide and protein
CC sequences from the Corynebacterium bacterium Corynebacterium glutamicum. These
CC are useful for identifying the mutation point of a gene derived from a
CC mutant of corynebacterium bacterium, measuring expression amount and analyzing
CC the expression profile or expression pattern of a gene derived from
CC Corynebacterium bacterium, and identifying a homologue of a gene derived from
CC corynebacterium bacterium. Corynebacterium bacteria are useful for producing amino
CC acids, nucleic acids, vitamins, saccharides and organic acids,
CC particularly L-lysine. The present sequence is a protein described in the
CC exemplification of the invention. Note: The sequence data for this patent
CC did not form part of the printed specification, but was obtained in
CC electronic format directly from the European Patent Office
XX
SQ Sequence 165 AA;
Query Match 57.0%; Score 473.5; DB 4; Length 165;
Best Local Similarity 57.3%; Pred. No. 5.3e-44;
Matches 94; Conservative 24; Mismatches 45; Indels 1; Gaps 1;
QY 1 MAQITLRGNAINTVGSIPAVGSPAPATLTGDLGVISSDQFRGKSVLTINIFPSVDTPVC 60
DB 1 MATHHFGGNETATSGELPQVGDMLAEFNLVNTLGRVSSXDFGRKVLINIFPSVDTPVC 60
QY 61 ATSVRTDERAAA-SGATVLCVSKDLPPAOKRFGCAGSTENWMPASAFRDSFGEDYGTI 119
DB 61 ATSVRKNEAAASLENTTVLCISKDLPPALGRFCAGIENVTVSFRSTFGEDNGIVL 120
QY 120 ADGPMAGLARAIVIGADGNVATTELVPBIAOBPNYEAALAL 163
DB 121 EGSPLKGLARSVIVVDENGVATQVDEISTEPDYDALAQL 164
RESULT 8
AAB78902
ID AAB78902 standard; protein; 165 AA.
XX
XX AAB78902;
XX
XX 30-APR-2001 (first entry)
XX
DB C. glutamicum SRT protein sequence SEQ ID NO:64.
XX
XX Corynebacterium glutamicum; stress; resistance; tolerance; SRT;
KW fine chemical production; organic acid; proteinogenic amino acid;
KW nonproteinogenic amino acid; purine base; pyrimidine base; nucleoside;
KW nucleoside; lipid; saturated fatty acid; unsaturated fatty acid; diol;
KW carbohydrate; aromatic compound; vitamin; cofactor; polyketide; enzyme;
KW evolutionary study; environmental hazard; fermentation.
XX
OS Corynebacterium glutamicum.
XX
XX WO200100804-A2.
XX
XX PD 04-JAN-2001.
XX
XX 23-JUN-2000; 2000WO-IB000922.
XX
XX 25-JUN-1999; 99US-0141031P.
PR 01-JUL-1999; 99DE-01030429.
XX 01-JUL-1999; 99US-0142692P.
PR 08-JUL-1999; 99DE-01031413.
PR 08-JUL-1999; 99DE-01031457.
PR 08-JUL-1999; 99DE-01031541.
PR 09-JUL-1999; 99DE-01032209.

PR 09-JUL-1999; 99DE-01032230.
PR 14-JUL-1999; 99DE-01032914.
PR 27-AUG-1999; 99DE-01040764.
PR 27-AUG-1999; 99US-0151214P.
PR 31-AUG-1999; 99DE-01041382.
XX
XX (BADI) BASF AG.
XX
XX Pompejus M, Kroeger B, Schroeder H, Zelder O, Habberhauer G;
PI Lee H, Kim H;
XX
XX WPI; 2001-061972/07.
DR N-PSDB; AAF71015.
XX
XX New isolated Corynebacterium glutamicum nucleic acid encoding a stress,
PT tolerance or resistance protein, for production or modulation of
PT production of fine chemicals, such as, e.g. amino acids, lipids,
PT carbohydrates, or enzymes.
XX
XX Claim 20; Page 211-212; 526bp; English.
XX
CC AAF70984 to AAF71133 encode the Corynebacterium glutamicum stress,
CC resistance, and tolerance (SRT) proteins given in AAB78871 to AAB79020.
CC The C. glutamicum SRT genes (I) can be used in vectors (II) for
CC expression in host cells and production of fine chemicals, such as, an
CC organic acid, a proteinogenic or nonproteinogenic amino acid (preferred),
CC a purine or pyrimidine base, a nucleoside, a nucleotide, a lipid, a
CC saturated or unsaturated fatty acid, a diol, a carbohydrate, an aromatic
CC compound, a vitamin, a cofactor, a polyketide, or an enzyme. The fine
CC chemical production can be modulated. The presence of (I) or the SRT
CC proteins (II) encoded by them are used for diagnosing the presence or
CC activity of Corynebacterium diptheriae. (I), (II), (III) and host cells
CC containing them can be used to map the genomes of organisms related to C.
CC glutamicum, to identify and localize C. glutamicum sequences of interest,
CC in evolutionary studies, in determination of SRT protein regions required
CC for function, in modulating the SRT protein activity, and in modulating
CC the activity of an SRT pathway. (II) are used to permit C. glutamicum to
CC survive in an environment that is normally environmentally or chemically
CC hazardous to it. (I) and protein molecules encoded by it increase the
CC survival of C. glutamicum to chemical and environmental hazards and
CC provide a means for continued growth and multiplication in large scale
CC fermentative growth conditions. By increasing the growth rate or
CC maintaining a normal growth rate in poor or toxic conditions, the yield,
CC production and/or efficiency or production of fine chemicals from a
CC culture may be increased
XX
XX Sequence 165 AA;
SQ
Query Match 56.1%; Score 466.5; DB 4; Length 165;
Best Local Similarity 56.7%; Pred. No. 3.1e-43;
Matches 93; Conservative 23; Mismatches 47; Indels 1; Gaps 1;
QY 1 MAQITLRGNAINTVGSIPAVGSPAPATLTGDLGVISSDQFRGKSVLTINIFPSVDTPVC 60
DB 1 MATHHFGGNETATSGELPQVGDMLAEFNLVNTLGRVSSXDFGRKVLINIFPSVDTPVC 60
QY 61 ATSVRTDERAAA-SGATVLCVSKDLPPAOKRFGCAGSTENWMPASAFRDSFGEDYGTI 119
DB 61 ATSVRKNEAAASLENTTVLCISKDLPPALGRFCAGIENVTVSFRSTFGEDNGIVL 120
QY 120 ADGPMAGLARAIVIGADGNVATTELVPBIAOBPNYEAALAL 163
DB 121 EGSPLKGLARSVIVVDENGVATQVDEISTEPDYDALAQL 164
RESULT 9
AAB79108
ID AAB79108 standard; protein; 165 AA.
XX
XX AAB79108;
XX
XX 30-APR-2001 (first entry)
XX

DE	Corynebacterium glutamicum HA protein sequence SEQ ID NO:172.
XX	
KW	Corynebacterium glutamicum; homeostasis; adaptation; HA protein;
KW	fine chemical production; organic acid; proteinogenic amino acid;
KW	nonproteinogenic amino acid; purine base; pyrimidine base; nucleoside;
KW	nucleoside; lipid; saturated fatty acid; unsaturated fatty acid; diol;
KW	carbohydrate; aromatic compound; vitamin; cofactor; polyketide; enzyme;
KW	diagnosis; Corynebacterium diptheriae; genetic engineering;
KW	Brevibacterium; environmental condition.
XX	
OS	Corynebacterium glutamicum.
XX	
PN	WO200100842-A2.
PD	
XX	
XX	04-JAN-2001.
PF	
XX	23-JUN-2000; 2000MO-IB000911.
XX	
PR	25-JUN-1999; 99US-0141031P.
PR	08-JUL-1999; 99DE-01031636.
PR	09-JUL-1999; 99DE-01032125.
PR	09-JUL-1999; 99DE-01032126.
PR	09-JUL-1999; 99DE-01032127.
PR	09-JUL-1999; 99DE-01032128.
PR	09-JUL-1999; 99DE-01032129.
PR	09-JUL-1999; 99DE-01032226.
PR	14-JUL-1999; 99DE-01032920.
PR	14-JUL-1999; 99DE-01032922.
PR	14-JUL-1999; 99DE-01032924.
PR	14-JUL-1999; 99DE-01032928.
PR	14-JUL-1999; 99DE-01032930.
PR	14-JUL-1999; 99DE-01032933.
PR	14-JUL-1999; 99DE-01032935.
PR	14-JUL-1999; 99DE-01032973.
PR	14-JUL-1999; 99DE-01033002.
PR	14-JUL-1999; 99DE-01033003.
PR	14-JUL-1999; 99DE-01033005.
PR	14-JUL-1999; 99DE-01033006.
PR	31-AUG-1999; 99DE-01041378.
PR	31-AUG-1999; 99DE-01041379.
PR	31-AUG-1999; 99DE-01041390.
PR	31-AUG-1999; 99DE-01041391.
PR	03-SEP-1999; 99DE-01042088.
XX	
PA	(BADI) BASF AG.
XX	
PI	Pompejus M, Kroeger B, Schroeder H, Zelder O, Habenhauer G;
XX	
DR	WPI; 2001-061974/07.
XX	
XX	N-PSDB; AAF71223.
PT	
PT	New isolated Corynebacterium glutamicum nucleic acid for production or
PT	modulation of production of fine chemicals such as amino acids,
PT	nucleosides, nucleotides, lipids, fatty acids, carbohydrates, vitamins or
PT	enzymes.
XX	
PS	Claim 20; Page 379-380; 712pp; English.
XX	
CC	AAE71138 to AAF71357 encode the Corynebacterium glutamicum homeostasis
CC	and adaptation (HA) proteins given in AAB79023 to AAB79242. The C.
CC	glutamicum HA genes (I) can be used in vectors for expression in host
CC	cells and production of fine chemicals, such as, an organic acid,
CC	proteinogenic or nonproteinogenic amino acid (preferred), purine or
CC	pyrimidine base, nucleoside, nucleotide, lipid, saturated or unsaturated
CC	fatty acid, diol, carbohydrate, aromatic compound, vitamin, cofactor,
CC	polyketide or enzyme. The amino acids produced can be lysine, glutamine,
CC	glutamate, alanine, aspartate, glycine, serine, threonine, methionine,
CC	cysteine, valine, leucine, isoleucine, arginine, proline, histidine,
CC	tyrosine, phenylalanine, or tryptophan. The fine chemical production can
CC	be modulated. The presence of (I) or HA proteins encoded by them are used
CC	for diagnosing the presence or activity of Corynebacterium diptheriae.
CC	(I) can be used to map the C. glutamicum genome or can be used as markers
CC	for genetically engineered Corynebacterium or Brevibacterium. The HA

CC glutamicum encoded by the (1) are used to maintain homeostasis in C.
CC glutanicum or help the microorganism to adapt to different environmental
CC conditions
XX
XX Sequence 165 AA,
SQ

Query Match 56.1%; Score 466.5; DB 4; Length 165;
Best Local Similarity 56.7%; Pred. No. 3,1e-43;
Matches 93; Conservative 23; Mismatches 47; Indels 1; Gaps 1;

OY 1 MAQITLKGMAINTVGELEPAVGSPAPAFTLTGGLGVISDQFRGKSVLLNIFPSVDTPVC 60
Db |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
1 MAKTHPGNEPATSTGELPQVGDNLAEFNLVNTLELGVSXKDFGRKLVLNFPSEDTGVC 60

OY 61 ATSVRTTDERAAA-SGATVLCVSKDLPPAQRFCAAGTERVMWASA.FRDSFGSDYGTVI 119
Db |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
61 ATSVRKTFEAAASILENTTVLCISXDLPFLALGRFCSAEGIEWTVFSARSTRFGBDNQIVL 120

OY 120 ADGSMAGLIARAIVIGADGNVAATELVPEIAOSEPNTEAALL 163
Db |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
121 EGSPILKGLAKRSVAVVBENKAAYTQLVDELTFFPYDALACL 164

RESULT 10
ADF04323
ID ADF04323 standard; protein; 172 AA.

XX ADF04323;
AC ADF04323;
XX
XX DT 12-FEB-2004 (first entry)
DB Bacterial polypeptide #436.

XX Proteus mirabilis infection; bacterial infection; antibacterial;
KW immunostimulant.
XX
XX OS Proteus mirabilis.
XX
XX PN US6605709-B1.
XX
XX PD 12-AUG-2003.
XX
XX PF 05-APR-2000; 2000US-00543681.
XX
XX PR 09-APR-1999; 99US-0128706P.
XX
XX PA (GENO-) GENOME THERAPEUTICS CORP.
XX
XX PI Breton GL;
XX
XX WIPI; 2003-895291/82.
DR N-PADB; ADF00151.
XX

PT New Proteus mirabilis polypeptides and polynucleotides, useful as
PT reagents for diagnosis of bacterial disease, as components of
PT antibacterial vaccines, as targets for antibacterial drugs, or as
PT biocontrol agents for plants.
XX
XX PS Disclosure; SEQ ID NO 4608; 870pp; English.

XX The invention relates to new Proteus mirabilis polypeptides and
XX polynucleotides. The invention also relates to antibodies against the
XX polypeptides, methods for producing the polypeptides, a method of
XX generating vaccines for immunising an individual against P. mirabilis, a
XX method for evaluating a compound for the ability to bind a P. mirabilis
XX polypeptide and a method for screening test compounds for anti-bacterial
XX activity. The polypeptides, preventing and polynucleotides are useful as molecular
XX targets for diagnosing, preventing and treating pathological conditions
XX resulting from bacterial infection, as reagents for diagnosis of
XX bacterial diseases, as components of antibacterial vaccines, as targets
XX for antibacterial drugs or as bio-control agents for plants. This
XX sequence represents a Proteus mirabilis polypeptide of the invention.

SQ Sequence 172 AA;
 Query Match 53.1%; Score 441; DB 7; Length 172;
 Best Local Similarity 55.2%; Pred. No. 2.3e-40;
 Matches 90; Conservative 21; Mismatches 50; Indels 2; Gaps 2;
 QY 3 QITLRGNAINTVGEIPAVGSPAPAFITLTGDLGVISDQFRGKSVLNTIPPSVDTPTCAT 62
 DB 9 QVTIQGNAVTLAGNFPVVGQKADPSLVGKDINVSILAQFAGKKVNLIPPSVDTGCAA 68
 QY 63 SVRFPDRAA-AAAGATVLCVSKDLPFAQKRCGAEGETENWPAFAFDS-FGEDYGV 120
 DB 69 SVRFQFNVAEINLNTVVLCTISADLPFAQKRCGAEGLDNTVLTSTMGAEFKENYGA 128
 QY 121 DGPAGGLAARAVIVIGADGVAATVPEIAQENYEAALAA 163
 DB 129 SGPLAGLTSRAVIVLDESNNVITQVLDEITTEPNDAALAVL 171
 RESULT 11
 ID ABO61199 standard; protein; 176 AA.
 XX ABO61199;
 AC ABO61199;
 XX 29-JUL-2004 (first entry)
 DT 29-JUL-2004 (first entry)
 XX Klebsiella pneumoniae polypeptide seqid 7716.
 DB Klebsiella pneumoniae polypeptide seqid 7716.
 KM Recombinant expression vector; transcription regulatory element;
 KW Klebsiella pneumoniae protein; antibacterial; Vaccine.
 XX Klebsiella pneumoniae.
 OS Klebsiella pneumoniae.
 XX US6610836-B1.
 PN 26-AUG-2003.
 PD 27-JAN-2000; 2000US-00489039.
 PF 29-JAN-1999; 99US-0117747P.
 PR (GENO-) GENOME THERAPEUTICS CORP.
 PA Breton GL, Osborne M;
 PI WPI; 2003-895346/82.
 DR N-PSDB; ACH94750.
 DR 29-JAN-1999; 99US-0117747P.
 XX New nucleic acid encoding a Klebsiella pneumoniae polypeptide, useful for
 PT preparing a vaccine composition against Klebsiella pneumoniae.
 CC Disclosure; SEQ ID NO 7716; 932bp; English.
 PS The invention describes a new isolated nucleic acid encoding a Klebsiella
 XX pneumoniae polypeptide. Also described are: a recombinant expression
 CC vector comprising the nucleic acid, operably linked to a transcription
 CC regulatory element; and a cell comprising the recombinant expression
 CC vector. The nucleic acid is useful for preparing a vaccine composition
 CC against Klebsiella pneumoniae. This is the amino acid sequence of a
 CC Klebsiella pneumoniae polypeptide of the invention
 XX
 SO Sequence 176 AA;
 Query Match 50.4%; Score 418.5; DB 7; Length 176;
 Best Local Similarity 53.0%; Pred. No. 7.3e-39;
 Matches 89; Conservative 24; Mismatches 52; Indels 3; Gaps 3;
 QY 1 MAQ-ITLRGNAINTVGEIPAVGSPAPAFITLTGDLGVISDQFRGKSVLNTIPPSVDTPT 59
 DB 9 MSQVHFGNPNVANSIPQAGSKAQFTLVADLSVITGQFAGKKVNLIPPSIDITGV 68
 QY 60 CATSVRTFDERAA-ASGATVLCVSKDLPFAQKRCGAEGETENWPAFAFDS-FGEDYGV 117
 DB 121 AIDGFLKGLAARAVIVIGADGVAATVPEIAQENYEAALAA 163

DB 69 CAASVRKFNQLAALDNTVVLCTISADLPFAQSRFCGAEGLSNVTLSTLRGASFLADYGV 128
 QY 118 TIADGPMAGLARAIVIVIGADGVAATVPEIAQENYEAALAA 165
 DB 129 AIAFGPLAGLAARAVIVIDENDQVLYSQLVNEITTEPNDAALAA 176
 RESULT 12
 ID AAU29352 standard; protein; 168 AA.
 XX AAU29352;
 AC AAU29352;
 XX 18-DEC-2001 (first entry)
 DT 18-DEC-2001 (first entry)
 DE Novel mar regulated protein (NMR) #24.
 KW mar regulated polypeptide; NMR; microbial infection; antibacterial.
 OS Becherichia coli.
 XX WO200170776-A2.
 PN 27-SEP-2001.
 PD 08-MAR-2001; 2001WO-US007478.
 PF 10-MAR-2000; 2000US-0186362P.
 PR (TUFT) TUFTS COLLEGE.
 PA Levy SB, Barbosa TM, Alekshun MN;
 PI WPI; 2001-602769/68.
 DR N-PSDB; AAS46252.
 DR Identifying compounds that modulate a newly identified mar regulated
 PT polypeptide activity, useful as antimicrobial compounds, involves
 PT contacting the polypeptide with a test compound.
 XX Disclosure; Page 309-310; 526pp; English.
 PS The invention relates to a method of identifying compounds that modulate
 XX a newly identified mar regulated (NMR) polypeptide activity. The method
 CC comprises contacting an NMR polypeptide with a test compound under
 CC interaction conditions, determining the ability of the compound to
 CC modulate the activity or expression of the polypeptide, and selecting the
 CC modulators. NMR nucleic acids and polypeptides are used in the treatment
 CC of microbial infections, and in screening for modulators of NMR
 CC expression and activity. These modulators can be used to reduce the
 CC infectivity of a microbe on a surface, and the virulence of a microbe in
 CC a subject suffering from an infection. AAU29329-AAU29379 represent
 CC Becherichia coli NMR amino acid sequences of the invention
 XX
 SO Sequence 168 AA;
 Query Match 50.2%; Score 417.5; DB 4; Length 168;
 Best Local Similarity 52.4%; Pred. No. 8.9e-38;
 Matches 88; Conservative 23; Mismatches 54; Indels 3; Gaps 3;
 QY 1 MAQ-ITLRGNAINTVGEIPAVGSPAPAFITLTGDLGVISDQFRGKSVLNTIPPSVDTPT 59
 DB 1 MSQVHFGNPNVANSIPQAGSKAQFTLVADLSVITGQFAGKKVNLIPPSIDITGV 60
 QY 60 CATSVRTFDERAA-ASGATVLCVSKDLPFAQSRFCGAEGETENWPAFAFDS-FGEDYGV 117
 DB 61 CAASVRKFNQLAALDNTVVLCTISADLPFAQSRFCGAEGLSNVTLSTFRNAEFLQAYGV 120
 QY 118 TIADGPMAGLARAIVIVIGADGVAATVPEIAQENYEAALAA 165
 DB 121 AIDGFLKGLAARAVIVIDENDQVLYSQLVNEITTEPNDAALAVL 168

Genomic sequence for *Listeria monocytogenes*, useful e.g. for treatment and prevention of *Listeria* and related bacterial infections, and related polypeptides.

Search completed: April 14, 2006, 17:18:22
Job time : 67.931 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using SW model

Run on: April 14, 2006, 17:19:08 ; Search time 9.67524 Seconds
(without alignments)
1640.866 Million cell updates/sec

Title: US-10-620-246-8

Perfect score: 831

Sequence: 1 MAQITLRGNAINTVGBLPANV.....LVPEIAQEPNYEAAALGA 165

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	831	100.0	165	2 H70635	thioredoxin peroxi
2	486	58.5	165	2 B83328	thiol peroxidase p
3	420	50.5	167	2 AG0285	thiol peroxidase (
4	417.5	50.2	168	2 B85752	thiol peroxidase (
5	417.5	50.2	168	2 JC5504	thioredoxin peroxi
6	417.5	50.2	168	2 G90866	thiol peroxidase (
7	405	48.7	165	2 G64090	thioredoxin peroxi
8	395.5	47.6	168	2 AG0659	thiol peroxidase (
9	352	42.4	164	2 C82274	thioredoxin peroxi
10	349	42.0	164	2 JC2570	probable thioredox
11	342.5	41.2	165	2 AH1635	thiol peroxidases
12	340.5	41.0	165	2 AG1272	thiol peroxidases
13	324.5	39.0	166	2 B84049	thiol peroxidase (
14	320.5	38.6	172	2 A95192	thioredoxin peroxi
15	320.5	38.6	172	2 F98058	thiol peroxidase (
16	314	37.8	175	2 C81349	thioredoxin peroxi
17	312	37.5	164	2 B89955	hypothetical prote
18	303.5	36.5	167	2 F69992	thioredoxin peroxi
19	303.5	36.5	169	2 H70343	thioredoxin peroxi
20	294.5	35.4	163	2 T11552	thioredoxin peroxi
21	294.5	35.4	163	2 B43583	thioredoxin peroxi
22	266.5	32.1	164	2 S61913	thioredoxin peroxi
23	259.5	31.2	164	2 C97306	thiol peroxidase,
24	254	30.6	162	2 D71861	thioredoxin peroxi
25	250	30.1	166	2 H86668	thiol peroxidase (
26	228.5	27.5	160	2 H86662	thiol peroxidase (
27	150	18.1	165	2 D90600	thiol peroxidase (
28	142	17.1	159	2 H75424	bacterioferritin c
29	138.5	16.7	164	2 H72518	probable bacteriof

30	135.5	16.3	156	2 G90434	hypothetical prote
31	133.5	16.1	157	2 E72332	thioredoxin peroxi
32	133	16.0	151	2 H75298	probable thiol-ope
33	128.5	15.5	155	2 AD2801	bacterioferritin c
34	128.5	15.5	155	2 F97580	bacterioferritin c
35	122.5	14.7	159	2 AH2118	bacterioferritin c
36	122.5	14.7	167	2 G70341	bacterioferritin c
37	122.5	14.7	156	2 T16005	hypothetical prote
38	121.5	14.6	257	2 J00064	MERS protein - mou
39	120.5	14.5	145	2 AH2102	hypothetical prote
40	118.5	14.3	157	2 F70870	probable bcp prote
41	118.5	14.3	195	2 JC2258	substrate protein
42	118	14.2	153	2 AH2203	bacterioferritin c
43	117.5	14.1	198	2 A57716	thiol-specific ant
44	116.5	14.0	185	2 AF0661	probable thiol per
45	116.5	14.0	200	2 S76284	hypothetical prote

ALIGNMENTS

RESULT 1

H70635
thioredoxin peroxidase (EC 1.11.1.-) - Mycobacterium tuberculosis (strain H37Rv)
N/Alternate names: scavengase; thiol peroxidase p20
C/Species: Mycobacterium tuberculosis
C/Date: 17-Jul-1998 #sequence_rev1510 17-Jul-1998 #text_change 09-Jul-2004
C/Accession: H70635
R/Colo, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, R.; Connor, R.; Davies, R.; Devlin, K.; Feldman, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S. Nature 393, 537-544, 1998
A/Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A/Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A/Reference number: A70500; MIMD:9825987; PMID:9634230
A/Accession: H70635
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-165 <COL>
A/Cross-references: UNIPROT:P95282; UNIPARC:UP1000004A4B5; GB:Z84498; GB:AL123456; NID:
A/Experimental source: strain H37Rv
C/Genetics:
A/Gene: tpx
C/Superfamily: thioredoxin peroxidase
C/Keywords: oxidoreductase; redox-active disulfide
F:60-93/Disulfide bonds: redox-active #status predicted

Query Match 100.0%; Score 831; DB 2; Length 165;
Best Local Similarity 100.0%; Pred. No. 3.5e-70;
Matches 165; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAQITLRGNAINTVGBLPANVATTLTGGLGVSSQPFKSKVTLNIFPVDPPVC 60
DB 1 MAQITLRGNAINTVGBLPANVSPAPATTLTGGLGVSSQPFKSKVTLNIFPVDPPVC 60

QY 61 ATSVRTPEBAASGATVLCVSKDLPAPQAFGAGETGVMPAPASFRDSFGSEYGTIA 120
DB 61 ATSVRTPEBAASGATVLCVSKDLPAPQAFGAGETGVMPAPASFRDSFGSEYGTIA 120

QY 121 DGPAGILAAIVIVIGADGNVATTELVPETIAQEPNYEAAALGA 165
DB 121 DGPAGILAAIVIVIGADGNVATTELVPETIAQEPNYEAAALGA 165

RESULT 2
E83328
thiol peroxidase P2532 [imported] - Pseudomonas aeruginosa (strain PAO1)
C/Species: Pseudomonas aeruginosa
C/Date: 15-Sep-2000 #sequence_rev1510 15-Sep-2000 #text_change 09-Jul-2004
C/Accession: E83328
R/Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Berman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim, J.; Lory, S.; Olson, M.V.

thioredoxin peroxidase (EC 1.11.1.-) [similarity] - Vibrio cholerae (strain N16961 serog C) Species: Vibrio cholerae
 C/Date: 18-Aug-2000 #sequence_rev1510n 20-Aug-2000 #text_change 09-Jul-2004
 C/Accession: C82374
 R:Heidelberg, J.F.; Eilen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R.J.; Chardon, D.; Ermolova, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragol, I.; Sellers, F. 1, R.R.; Mekalanos, J.J.; Venter, J.C.; Frazer, C.M.
 Nature 406, 477-483, 2000
 A/Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
 A/Reference number: AB2035; PMID:20406833; PMID:10952301
 A/Accession: C82274
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-164 <HEI>
 A/Cross-references: UNIPROT:P39167; UNIPARC:UPI000000A88; GB:AE004168; GB:AE003852; NIT A/Experimental source: serogroup O1; strain N16961, biotype El Tor
 C/Genetics:
 A/Gene: VC0824
 A/Map position: 1
 C/Superfamily: thioredoxin peroxidase
 C/Keywords: oxidoreductase; redox-active disulfide
 F/59-93/Disulfide bonds: redox-active #status predicted

Query Match 42.4%; Score 352; DB 2; Length 164;

Best Local Similarity 43.2%; Pred. No. 1.4e-25;

Matches 70; Conservative 30; Mismatches 60; Indels 2; Gaps 2;

4 ITLRGNAINTVGBLPAVGPAPAFITLGTGDLGVISDQFGKSVLNTIPSPVTPVCATS 63

3 VTFQNNPVISISGSPFKVGRPLPSFTLCGADLNDLNEDFGKCIWISFISITPVCSSK 62

64 VRTFDRRAAASGATV-LCVSKDLPFAOKRFGAGETENWPAAPRD-SFGEDYGVTTAD 121

63 VKTLQNALMTNRDITVLLCVSADLPFAMSRCTEHAANVTNNSFFRPPATERRFGVNLNE 122

122 GPWAGLLARAIVIVIGADGNVATYELVPEIAQEPNYEALAAAL 163

123 GALRGLAARAVIVADEFGVITSHSELVNEITNEBDYRILMSL 164

Db

Query

123 GALRGLAARAVIVADEFGVITSHSELVNEITNEBDYRILMSL 164

123 GALRGLAARAVIVADEFGVITSHSELVNEITNEBDYRILMSL 164

123 GALRGLAARAVIVADEFGVITSHSELVNEITNEBDYRILMSL 164

123 GALRGLAARAVIVADEFGVITSHSELVNEITNEBDYRILMSL 164

123 GALRGLAARAVIVADEFGVITSHSELVNEITNEBDYRILMSL 164

123 GALRGLAARAVIVADEFGVITSHSELVNEITNEBDYRILMSL 164

123 GALRGLAARAVIVADEFGVITSHSELVNEITNEBDYRILMSL 164

123 GALRGLAARAVIVADEFGVITSHSELVNEITNEBDYRILMSL 164

123 GALRGLAARAVIVADEFGVITSHSELVNEITNEBDYRILMSL 164

123 GALRGLAARAVIVADEFGVITSHSELVNEITNEBDYRILMSL 164

123 GALRGLAARAVIVADEFGVITSHSELVNEITNEBDYRILMSL 164

123 GALRGLAARAVIVADEFGVITSHSELVNEITNEBDYRILMSL 164

123 GALRGLAARAVIVADEFGVITSHSELVNEITNEBDYRILMSL 164

123 GALRGLAARAVIVADEFGVITSHSELVNEITNEBDYRILMSL 164

123 GALRGLAARAVIVADEFGVITSHSELVNEITNEBDYRILMSL 164

123 GALRGLAARAVIVADEFGVITSHSELVNEITNEBDYRILMSL 164

123 GALRGLAARAVIVADEFGVITSHSELVNEITNEBDYRILMSL 164

123 GALRGLAARAVIVADEFGVITSHSELVNEITNEBDYRILMSL 164

123 GALRGLAARAVIVADEFGVITSHSELVNEITNEBDYRILMSL 164

123 GALRGLAARAVIVADEFGVITSHSELVNEITNEBDYRILMSL 164

123 GALRGLAARAVIVADEFGVITSHSELVNEITNEBDYRILMSL 164

123 GALRGLAARAVIVADEFGVITSHSELVNEITNEBDYRILMSL 164

123 GALRGLAARAVIVADEFGVITSHSELVNEITNEBDYRILMSL 164

123 GALRGLAARAVIVADEFGVITSHSELVNEITNEBDYRILMSL 164

123 GALRGLAARAVIVADEFGVITSHSELVNEITNEBDYRILMSL 164

123 GALRGLAARAVIVADEFGVITSHSELVNEITNEBDYRILMSL 164

123 GALRGLAARAVIVADEFGVITSHSELVNEITNEBDYRILMSL 164

123 GALRGLAARAVIVADEFGVITSHSELVNEITNEBDYRILMSL 164

123 GALRGLAARAVIVADEFGVITSHSELVNEITNEBDYRILMSL 164

123 GALRGLAARAVIVADEFGVITSHSELVNEITNEBDYRILMSL 164

123 GALRGLAARAVIVADEFGVITSHSELVNEITNEBDYRILMSL 164

123 GALRGLAARAVIVADEFGVITSHSELVNEITNEBDYRILMSL 164

123 GALRGLAARAVIVADEFGVITSHSELVNEITNEBDYRILMSL 164

123 GALRGLAARAVIVADEFGVITSHSELVNEITNEBDYRILMSL 164

thioredoxin peroxidase (EC 1.11.1.-) [similarity] - Vibrio cholerae (strain N16961 serog C) Species: Vibrio cholerae
 C/Date: 18-Aug-2000 #sequence_rev1510n 20-Aug-2000 #text_change 09-Jul-2004
 C/Accession: C82374
 R:Heidelberg, J.F.; Eilen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R.J.; Chardon, D.; Ermolova, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragol, I.; Sellers, F. 1, R.R.; Mekalanos, J.J.; Venter, J.C.; Frazer, C.M.
 Nature 406, 477-483, 2000
 A/Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
 A/Reference number: AB2035; PMID:20406833; PMID:10952301
 A/Accession: C82274
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-164 <HEI>
 A/Cross-references: UNIPROT:P39167; UNIPARC:UPI000000A88; GB:AE004168; GB:AE003852; NIT A/Experimental source: serogroup O1; strain N16961, biotype El Tor
 C/Genetics:
 A/Gene: VC0824
 A/Map position: 1
 C/Superfamily: thioredoxin peroxidase
 C/Keywords: oxidoreductase; redox-active disulfide
 F/59-93/Disulfide bonds: redox-active #status predicted

Query Match 42.4%; Score 352; DB 2; Length 164;

Best Local Similarity 43.2%; Pred. No. 1.4e-25;

Matches 70; Conservative 30; Mismatches 60; Indels 2; Gaps 2;

4 ITLRGNAINTVGBLPAVGPAPAFITLGTGDLGVISDQFGKSVLNTIPSPVTPVCATS 63

3 VTFQNNPVISISGSPFKVGRPLPSFTLCGADLNDLNEDFGKCIWISFISITPVCSSK 62

64 VRTFDRRAAASGATV-LCVSKDLPFAOKRFGAGETENWPAAPRD-SFGEDYGVTTAD 121

63 VKTLQNALMTNRDITVLLCVSADLPFAMSRCTEHAANVTNNSFFRPPATERRFGVNLNE 122

122 GPWAGLLARAIVIVIGADGNVATYELVPEIAQEPNYEALAAAL 163

123 GALRGLAARAVIVADEFGVITSHSELVNEITNEBDYRILMSL 164

Db

Query

123 GALRGLAARAVIVADEFGVITSHSELVNEITNEBDYRILMSL 164

123 GALRGLAARAVIVADEFGVITSHSELVNEITNEBDYRILMSL 164

123 GALRGLAARAVIVADEFGVITSHSELVNEITNEBDYRILMSL 164

123 GALRGLAARAVIVADEFGVITSHSELVNEITNEBDYRILMSL 164

123 GALRGLAARAVIVADEFGVITSHSELVNEITNEBDYRILMSL 164

123 GALRGLAARAVIVADEFGVITSHSELVNEITNEBDYRILMSL 164

123 GALRGLAARAVIVADEFGVITSHSELVNEITNEBDYRILMSL 164

123 GALRGLAARAVIVADEFGVITSHSELVNEITNEBDYRILMSL 164

123 GALRGLAARAVIVADEFGVITSHSELVNEITNEBDYRILMSL 164

123 GALRGLAARAVIVADEFGVITSHSELVNEITNEBDYRILMSL 164

123 GALRGLAARAVIVADEFGVITSHSELVNEITNEBDYRILMSL 164

123 GALRGLAARAVIVADEFGVITSHSELVNEITNEBDYRILMSL 164

123 GALRGLAARAVIVADEFGVITSHSELVNEITNEBDYRILMSL 164

123 GALRGLAARAVIVADEFGVITSHSELVNEITNEBDYRILMSL 164

123 GALRGLAARAVIVADEFGVITSHSELVNEITNEBDYRILMSL 164

123 GALRGLAARAVIVADEFGVITSHSELVNEITNEBDYRILMSL 164

123 GALRGLAARAVIVADEFGVITSHSELVNEITNEBDYRILMSL 164

123 GALRGLAARAVIVADEFGVITSHSELVNEITNEBDYRILMSL 164

123 GALRGLAARAVIVADEFGVITSHSELVNEITNEBDYRILMSL 164

123 GALRGLAARAVIVADEFGVITSHSELVNEITNEBDYRILMSL 164

123 GALRGLAARAVIVADEFGVITSHSELVNEITNEBDYRILMSL 164

123 GALRGLAARAVIVADEFGVITSHSELVNEITNEBDYRILMSL 164

123 GALRGLAARAVIVADEFGVITSHSELVNEITNEBDYRILMSL 164

123 GALRGLAARAVIVADEFGVITSHSELVNEITNEBDYRILMSL 164

123 GALRGLAARAVIVADEFGVITSHSELVNEITNEBDYRILMSL 164

123 GALRGLAARAVIVADEFGVITSHSELVNEITNEBDYRILMSL 164

123 GALRGLAARAVIVADEFGVITSHSELVNEITNEBDYRILMSL 164

123 GALRGLAARAVIVADEFGVITSHSELVNEITNEBDYRILMSL 164

123 GALRGLAARAVIVADEFGVITSHSELVNEITNEBDYRILMSL 164

123 GALRGLAARAVIVADEFGVITSHSELVNEITNEBDYRILMSL 164

123 GALRGLAARAVIVADEFGVITSHSELVNEITNEBDYRILMSL 164

123 GALRGLAARAVIVADEFGVITSHSELVNEITNEBDYRILMSL 164

123 GALRGLAARAVIVADEFGVITSHSELVNEITNEBDYRILMSL 164

123 GALRGLAARAVIVADEFGVITSHSELVNEITNEBDYRILMSL 164

A;Cross-references: UNIPROT:Q8YCU8; UNIPARC:UPI0000054F8B; GB:NC_003210; PLIN:CAC99661.1
A;Experimental source: strain KGD-e
C;Genetics:
A;Gene: lmo1583
C;Superfamily: thioredoxin peroxidase

Query Match	41.0%;	Score 340.5;	DB 2;	Length 165;
Best Local Similarity	44.5%;	Pred. No. 1.7e-24;		
Matches 73;	Conservative 27;	Mismatches 59;	Indels 5;	Gaps 3;

```

QY      1 MAQITLRGAIWTVGELPAVGSBPAPFTLTGGLGYLSIDQFGKSVTLNIPSPVDYPPVC 60
DQ      1 MTQVTFKANPVTLVGIRKRVGDRAPFTVVRNRLBSRYTLHDYDGKRLISVPSIDHSVC 60
QY      61 ATSVPTPERAA--ASGATVLCVSKDLPFAOKRPGCAEGTENWMPASAFRD--SEGDYGYT 118
DQ      61 STQTRKFFBKASNLNDTIVVLTISVDLPFAOKKCAABGLPNAITLSDHRDLSFGAEYGV 120
QY      119 IADGPMAGILRAIVITGADGWAAYTELYPEIIOBPDEYEAALAA 162
DQ      121 MKG---LRLLARSVFVWNAVGRIYVTEVVPBGSDHNPYERAIATA 161

```

RESULT 13
 B84049
 Chloredoxin peroxidase BH3194 [imported] - *Bacillus halodurans* (strain C-125)
 C/Species: *Bacillus halodurans*
 C/Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004
 C/Accession: B84049
 R/Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hirose
 Nucleic Acids Res. 28, 4317-4331, 2000
 A/Title: Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* and
 /Reference number: AB3650; MUID:20512582; PMID:11058312

A/Status: Preliminary
A/Molecule type: DNA
A/Residues: 1-166 <STO>
A/Cross-references: UNIPROT:Q9K13; UNIPARC:UPI000013728F; GB:AP001518; GB:BA000004; NIT
A/Experimental source: strain C-125
C/Genetics:
A/Gene: BH3194
C/Superfamily: thioredoxin peroxidase

Query Match 39.0%; Score 324.5; DB 2; Length 166;
Best Local Similarity 44.9%; Pred. No. 5.3e-23;
Matches 75; Conservative 19; Mismatches 68; Indels 5; Gaps 3;

RESULT 14
A95192
chial peroxidase [imported] - Streptococcus pneumoniae (strain TIGR4)
C/Species: Streptococcus pneumoniae
C/Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 24-Aug-2001
C/Accession: A95192
R/Tetcelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heid
on, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radume, D.; Holtzapfel,
nson, T.; Hickey, E.K.; Holt, I.E.
Science 293, 498-506, 2001
A/Authors: Lofcus, B.J.; Yang, P.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,
A./Title: Complete Genome Sequence of a Virulent Isolate of Streptococcus pneumoniae.
A./Reference number: A95000; MUID:21357209; PMID:11463916

A:Accession: A95192
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-172 <KOR>
A:Cross-references: UNIPARC:UPI0000165A91, GB:AE005672, P1DN:AAK75730.1, PID:g14973141,
A:Experimental source: strain TIGR4
C:Genetics:
A:Gene: SP1651
C:Superfamily: thioredoxin peroxidase

	Query Match	38.6%;	Score 320.5;	DB 2;	Length 172;	
	Best Local Similarity	44.5%;	Pred. No. 1.3e-22;			
	Matches	73;	Conservative	23;	Mismatches	63;
					Indels	5;
					Gaps	3
<hr/>						
Qy	4	ITLRGNAINTVGBLPAWGSPAPAFPTLLNGDGLGVSSDDPGRKSVLIANTFPSSVDTPVCAT.S	63			
Db	11	VTFPGNPVPSFPGKQLQWDRKALDFSLATTDTLSKKSLADFDGKKRKLVSVPISIDTGICSTQ	70			
<hr/>						
Qy	64	VRTEDER-AAASGAKTVLCYSKDLPFAQRFCGAETEN-VMPASA.FRDSFGSDYGTIAD	121			
Db	71	TTRNEELAGLDNTVVTLVSNMLPPAQRWCAGEIGDAIMTSDYFDSHSGRDYALLINE	130			
<hr/>						
Qy	122	GPMAGLLARAVIVTIGADGNVAYTELVPEPIAOBPNTVEALLAAAGA	165			
Db	131	W--H-LARRAVFLDVTNLTIRYEVVDNIINSEPNFEAIIAAAKA	171			

RESULT 15
P98058
thioredoxinlinked thiol peroxidase [EC:1.11.1.-] (imported) - Streptococcus pneumoniae (C)Species: Streptococcus pneumoniae
C/Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 02-Nov-2001
C/Accession: F98058
R/Hoskins, J.A.; Albom Jr., W.; Arnold, J.; Bursgett, S.; Dehoff, B.S.; e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; Mchren, S.; y, P.; Sun, P.M.; Winkler, M.E.
Y. Bacteriol. 183, 5709-5717, 2001
A/Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaekunas, S.R.
A/Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.
A/Reference number: A97872; MUID:21429245; PMID:11544234
A/Accession: F98058
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-172 <KOR>
C/Cross-references: UNIPARC:UP10000165A91, GB:AE007317, PIDN:AAL00299.1, PID:g15459155,
A/Gene: tpx
C/Superfamily: thioredoxin peroxidase
/Keywords: oxidoreductase

38.6%;	Score 320.5; DB 2; Length 172;
Query Match	44.5%; Pred. No. 1.3e-22;
Best Local Similarity	Matches 73; Conservative 23; Mismatches 63; Indels 5; Gaps 3
QY	4 ITTNGAINTVIGELPAVGSPPAFTLTGGDVGIVISSDQFRGKSVLLNTPPSVDTPVCATS 63
DB	11 VTFIENGVPVSFTGKGLQVGDKALDPSLFTTDLSSKSLADFDGKKVLSVPSIDTGHCSTQ 70
QY	64 VRTYDER-AAAASGATVLCVSKQLPPAQKRFEGARSTEN-WNPASAFRDSFGSDYGYTIND 121
DB	71 TRRNEETIAGIDNTVTVLVTSMDLPPAQKRWGARGLDMAIMLSDYFDSHSGFDYVALLINE 130
QY	122 GPMAIGLAPAIIVTVGADGNVATLEIPEIACQEPYVBAALLALGA 165
DB	131 W---HLIARAVFVLDITNTITIRYEVVDNINSEEPFEALIAAKA 171

Search completed: April 14, 2006, 17:34:20
Job time : 10.6752 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 14, 2006, 17:04:10 ; Search time 60.8796 Seconds

(without alignments)
1912.171 Million cell updates/sec

Title: US-10-620-246-8

Perfect score: 831
Sequence: 1 MAQITRGNAINTVGEIPAV.....LVPBIAQEPYEAALALGA 165

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Litering first 45 summaries

Database : UniProt_05.80.*
1: uniProt_sprot.*
2: uniProt_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	831	100.0	165	1 TPX_MYCBO	P66953 mycobacteri
2	831	100.0	165	1 TPX_MYCBO	P66952 mycobacteri
3	831	100.0	165	2 Q7AUD0_MYCBO	Q7AUD0 mycobacteri
4	705	84.8	164	2 Q8GF25_MYCPA	Q8GF25 mycobacteri
5	701	84.4	160	2 Q8AG72_MYCAV	Q8AG72 mycobacteri
6	519.5	62.5	167	2 Q6AA71_PROAC	Q6AA71 propionibac
7	504	60.6	165	2 Q7N201_CHRVO	Q7N201 chrysochlor
8	499	58.0	168	1 TPX_CHLRE	Q8K65 chlorobact
9	486	58.5	165	1 TPX_PSEBR	P57668 pseudomonas
10	477.5	57.5	179	2 Q7M37_PORGI	Q7M37 porphyromon
11	476	57.3	166	2 Q4KCR9_PSEBS	Q4KCR9 pseudomonas
12	473.5	57.0	165	1 TPX_CORGL	Q8H93 corynebacte
13	467	56.2	166	2 Q5LBP6_BACFN	Q5LBP6 bacteroides
14	467	56.2	166	2 Q6AS16_BACFN	Q6AS16 bacteroides
15	465	56.0	167	2 Q6AM98_DSPTS	Q6AM98 desulfocale
16	462	55.6	187	2 Q4NM07_9DELT	Q4NM07 anaeromyxob
17	458	55.1	166	2 Q881X9_PSEBS	Q881X9 pseudomonas
18	454	54.6	166	2 Q4ZTKE_PSSBY	Q4ZTKE pseudomonas
19	454	54.6	167	2 Q8A845_BACTN	Q8A845 bacteroides
20	453	54.5	166	2 Q88GY0_PSEBK	Q88GY0 pseudomonas
21	449.5	54.1	165	1 TPX_CORBP	Q8FG8 corynebacte
22	447	53.8	166	1 TPX_RALSO	Q8X90 ralestonia s
23	444.5	53.5	165	2 Q6MO7_BDBEA	Q6MO7 bdellovibri
24	420	50.5	167	1 TPX_YERPS	Q8Z42 yerersinia pe
25	420	50.5	167	2 Q6A71_YERPS	Q6A71 yerersinia ps
26	418.5	50.4	167	2 Q62GX0_BURMA	Q62GX0 burkholderi
27	418.5	50.4	167	2 Q63ON7_BURPS	Q63ON7 burkholderi
28	417	50.2	167	1 TPX_BCO57	P0A664 escherichia
29	417	50.2	167	1 TPX_BCO16	P0A883 escherichia
30	417	50.2	167	1 TPX_BCO11	P0A862 escherichia
31	417	50.2	167	1 TPX_SHIDY	P0A866 shigella dy

32	417	50.2	167	1 TPX_SHITL	P0A865 shigella fl
33	413.5	49.8	165	2 Q6NGY6_CORDI	Q6NGY6 corynebacte
34	408	49.1	165	1 TPX_PASMU	P57880 pasteurella
35	408	49.1	165	2 Q4QME7_HAB18	Q4QME7 haemophilus
36	406.5	48.9	167	2 Q6DS01_ERWCT	Q6DS01 erwina car
37	406	48.9	167	2 Q7N3Y3_PHOIL	Q7N3Y3 photorhabdu
38	405	48.7	165	1 TPX_HABIN	Q57349 haemophilus
39	402.5	48.4	167	2 Q6H800_BURML	Q6H800 burkholderi
40	402.5	48.4	168	1 TPX_SALTY	Q8ZP65 salmonella
41	402.5	48.4	168	2 Q57NY0_SALCH	Q57NY0 salmonella
42	402.5	48.4	168	2 Q5PHP7_SALPA	Q5PHP7 salmonella
43	400.5	48.2	167	2 Q4LIZ2_BURKH	Q4LIZ2 burkholderi
44	395.5	47.6	168	1 TPX_SALTY	Q8Z748 salmonella
45	379	45.6	166	2 Q5PAJ4_AZOSB	Q5PAJ4 azoarcus sp

ALIGNMENTS

RESULT 1
TPX_MYCBO STANDARD; PRT; 165 AA.

AC P66953; P95282;
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Probable thiol peroxidase (EC 1.11.1.-).
GN Name=tpx; OrderedLocustNames=Wb1967;
OS Mycobacterium bovis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium;
OC Mycobacterium tuberculosis complex.
OX NCBI_Taxid=1765;
RN [1]

RP NUCLEOTIDE SEQUENCE (LARGE SCALE GENOMIC DNA).

RC STRAIN=AF2122/97; PubMed=12788972; DOI=10.1073/pnas.1130426100;

RX MEDLINE=22709107; PubMed=12788972; DOI=10.1073/pnas.1130426100;

RA Garnier T., Biglmeier K., Camus J.-C., Medina N., Mansoor H.,

RA Pryor M., Dutfoy S., Grondin S., Lacroix C., Monsemp C., Simon S.,

RA Harris B., Atkins R., Doggett J., Mayes R., Keating L., Wheeler P.R.,

RA Parkhill J., Barrell B.G., Cole S.T., Gordon S.V., Hewinson R.G.,

RT "The complete genome sequence of Mycobacterium bovis.",

RL Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882(2003).

CC - FUNCTION: Has antioxidant activity. Could remove peroxides or

CC H(2)O(2) (By similarity).

CC - SIMILARITY: Belongs to the ahpc/TSA family. Tpx subfamily.

CC This Swiss-Prot entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use as long as its content is in no way modified and this statement is not

CC removed.

CC EMBL; BX248340; CAD94669.1; -; Genomic DNA.

DR HSSP; P72500; 1PSQ.

DR SMR; P66953; 2-165.

DR HAMAP; MF_00269; -; 1.

DR InterPro; IPR000866; Ahpc-TSA.

DR InterPro; IPR012336; Thiorodoxin-like.

DR InterPro; IPR012335; Thiorodoxin-fold.

DR PANTHER; PTHR10681:SP4; TPX; 1.

DR Pfam; PF00578; Ahpc-TSA; 1.

DR PROSITE; PS01265; TPX; 1.

KW Antioxidant; Complete proteome; Oxidoreductase; Peroxidase.

SQ SSQUNCE 165 AA; 16896 MW; B251D5DD8286F8 CRC64;

Query Match 100.0%; Score 831; DB 1; Length 165;

Best local similarity 100.0%; Pred. No. 6-5e-64;

Matches 165; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAQITRGNAINTVGEIPAVFTLTGGLGVYSSDQFRGKSVLLNTPSVDTPTVC 60

```

Db      1 MAQITLRGNAINTVGBLPAVGSPPAPFTLTGDLGVISSDQFRGKSVLTNIIPSVDTPMC 60
QY      61 ATSVRTFDEBPAASGATVLCVSKDLPFAOKRFGAGSTENVMPSAFRDSFGEDYGTIA 120
        |||
Db      61 ATSVRTFDEBPAASGATVLCVSKDLPFAOKRFGAGSTENVMPSAFRDSFGEDYGTIA 120
QY      121 DGPMAGLARAIVVIGADGNVATTELVPBIAOBPNTEAALALGA 165
        |||
        121 DGPMAGLARAIVVIGADGNVATTELVPBIAOBPNTEAALALGA 165
Db      121 DGPMAGLARAIVVIGADGNVATTELVPBIAOBPNTEAALALGA 165

RESULT 2
TPX_MYCTU STANDARD; PRT; 165 AA.
AC      P65952; P95282;
DT      01-NOV-1997 (Rel. 35, Created)
DT      25-OCT-2004 (Rel. 45, Last sequence update)
DE      10-MAY-2005 (Rel. 47, Last annotation update)
DE      Probable thiol peroxidase (EC 1.11.1.-).
GN      Name:tpx; OrderedLocuNames=Rv1932, MT1982; ORFNames=MTCY09F9.32c;
OS      Mycobacterium tuberculosis.
OC      Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC      Corynebacterineae; Mycobacteriaceae; Mycobacterium;
OC      Mycobacterium tuberculosis complex.
OX      NCBI_TaxID=1773;
[1]
RN      NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RP      STRAIN=H37Rv;
RX      MEDLINE=98295987; PubMed=9634230; DOI=10.1038/31159;
        Cole S.T., Broesch R., Parkhill J., Garnier T., Churcher C.M.,
        Harris D.E., Gordon S.V., Eigmeier K., Gas S., Barry C.B. III,
        Tekala P., Badcock K., Baaham D., Brown D., Chillingworth T.,
        Connor R., Davies R.M., Devlin K., Felwell T., Gentles S., Hamlin N.,
        Holroyd S., Hornsby T., Jagels K., Krogh A., McLean M.A.,
        Murphy L.D., Oliver S., Osborne J., Quail M.A., Rajandream M.A.,
        Rogers J., Rutter S., Seeger K., Skelton S., Squares R.,
        Sultson J.E., Taylor K., Whitehead S., Barrrell B.G.;
        "Deciphering the biology of Mycobacterium tuberculosis from the
        RT complete genome sequence.";
        Nature 393:537-544(1998).
[2]
RN      NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RP      STRAIN=CDC 1551 / Oshkosh;
RX      MEDLINE=2206494; PubMed=12218036;
        DOI=10.1126/JB.184.19.5479-5490.2002;
        Rietichmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
        Peterson J.D., Deboy R.T., Dodson R.J., Gwinn M.L., Haft D.H.,
        Hickey E.K., Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D.,
        Salzberg S.L., Delcher A., Utterback T.R., Weidman J.F., Kouri H.M.,
        Gali J., Mikula A., Bishai W., Jacobs W.R. Jr., Venter J.C.,
        Fraser C.M.;
        "Whole-genome comparison of Mycobacterium tuberculosis clinical and
        laboratory strains.";
        J. Bacteriol. 184:5479-5490(2002).
-1- FUNCTION: Has antioxidant activity. Could remove peroxides or
        H(2)O(2) (By similarity).
-1- SIMILARITY: Belongs to the ahpc/TSA family. Tpx subfamily.
CC      This Swiss-Prot entry is copyright. It is produced through a collaboration
        between the Swiss Institute of Bioinformatics and the EMBL Outstation -
        the European Bioinformatics Institute. There are no restrictions on its
        use as long as its content is in no way modified and this statement is not
        removed.
CC      -----
DR      EMBL; BX842578; CAB06529.1; -; Genomic_DNA.
DR      EMBL; AB000516; AAK46254.1; -; Genomic_DNA.
DR      PIR; H70635; H70635.
DR      PDB; 1XVO; X-ray; A=1-165.
DR      SMR; P65852; 2-165.
DR      TIGR; MT1982; -.
DR      Tuberculist; Rv1932; -.
DR      HAMAP; MF_00269; -; 1.
DR      InterPro; IPR00866; Ahpc-TSA.

```

```

DR      InterPro; IPR012336; Thiorodoxin-like.
DR      InterPro; IPR012335; Thiorodoxin_fold.
DR      InterPro; IPR002065; Tpx.
DR      PANTHER; PTHR10681; SFA; Tpx; 1.
DR      Pfam; PF00578; Ahpc-TSA; 1.
DR      PROSITE; PS01265; Tpx; 1.
DR      3D-structure; Antioxidant; Complete proteome; Oxidoreductase;
        Peroxidase.
SQ      SEQUENCE 165 AA; 16896 MW; B251D5DADE2286FE CRC64;

Query Match      100.0%; Score 831; DB 1; Length 165;
Best Local Similarity 100.0%; Pred. No. 6-5e-64;
Matches 165; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MAQITLRGNAINTVGBLPAVGSPPAPFTLTGDLGVISSDQFRGKSVLTNIIPSVDTPMC 60
        |||
Db      1 MAQITLRGNAINTVGBLPAVGSPPAPFTLTGDLGVISSDQFRGKSVLTNIIPSVDTPMC 60
QY      61 ATSVRTFDEBPAASGATVLCVSKDLPFAOKRFGAGSTENVMPSAFRDSFGEDYGTIA 120
        |||
Db      61 ATSVRTFDEBPAASGATVLCVSKDLPFAOKRFGAGSTENVMPSAFRDSFGEDYGTIA 120
QY      121 DGPMAGLARAIVVIGADGNVATTELVPBIAOBPNTEAALALGA 165
        |||
        121 DGPMAGLARAIVVIGADGNVATTELVPBIAOBPNTEAALALGA 165
Db      121 DGPMAGLARAIVVIGADGNVATTELVPBIAOBPNTEAALALGA 165

RESULT 3
Q7ANI0 MYCBO
ID      Q7ANI0_MYCBO PRELIMINARY; PRT; 165 AA.
AC      Q7ANI0;
DT      05-JUL-2004 (TrEMBLrel. 27, Created)
DT      05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT      05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE      Tpx protein.
OS      Mycobacterium bovis.
GN      Mycobacterium bovis.
OC      Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC      Corynebacterineae; Mycobacteriaceae; Mycobacterium;
OC      Mycobacterium tuberculosis complex.
OX      NCBI_TaxID=1765;
[1]
RN      NUCLEOTIDE SEQUENCE.
RP      Inwald J.K.;
RL      Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR      EMBL; A1486956; CAD31796.1; -; Genomic_DNA.
DR      SR; Q7ANI0.2-165.
DR      GO; GO:0009031; F:thiol peroxidase activity; IEA.
DR      InterPro; IPR000866; Ahpc-TSA.
DR      InterPro; IPR012336; Thiorodoxin-like.
DR      InterPro; IPR002065; Tpx.
DR      Pfam; PF00578; Ahpc-TSA; 1.
DR      PROSITE; PS01265; Tpx; 1.
SQ      SEQUENCE 165 AA; 16896 MW; B251D5DADE2286FE CRC64;

Query Match      100.0%; Score 831; DB 2; Length 165;
Best Local Similarity 100.0%; Pred. No. 6-5e-64;
Matches 165; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MAQITLRGNAINTVGBLPAVGSPPAPFTLTGDLGVISSDQFRGKSVLTNIIPSVDTPMC 60
        |||
Db      1 MAQITLRGNAINTVGBLPAVGSPPAPFTLTGDLGVISSDQFRGKSVLTNIIPSVDTPMC 60
QY      61 ATSVRTFDEBPAASGATVLCVSKDLPFAOKRFGAGSTENVMPSAFRDSFGEDYGTIA 120
        |||
Db      61 ATSVRTFDEBPAASGATVLCVSKDLPFAOKRFGAGSTENVMPSAFRDSFGEDYGTIA 120
QY      121 DGPMAGLARAIVVIGADGNVATTELVPBIAOBPNTEAALALGA 165
        |||
        121 DGPMAGLARAIVVIGADGNVATTELVPBIAOBPNTEAALALGA 165
Db      121 DGPMAGLARAIVVIGADGNVATTELVPBIAOBPNTEAALALGA 165

RESULT 4

```

08GP25 MYCPA PRELIMINARY; PRT; 164 AA.

ID 08GP25 MYCPA PRELIMINARY; PRT; 164 AA.

AC 08GP25_Q73ZP0; Created

DT 01-MAR-2003 (TRMBLrel. 23, Last sequence update)

DT 01-MAR-2003 (TRMBLrel. 23, Last sequence update)

DT 01-FEB-2005 (TRMBLrel. 29, Last annotation update)

DE 20 kDa protein (Tpx).

GN Name:tpx; OrderedLocustNames=MA1653;

OS Mycobacterium paratuberculosis.

OC Bacteria; Actinobacteria; Actinomycetales;

OC Corynebacteriaceae; Mycobacteriaceae; Mycobacterium;

OC Mycobacterium avium complex (MAC).

OX NCBI_TaxID=1770;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RX MEDLINE=22511501; PubMed=12623276; DOI=10.1016/S0882-4010(02)00209-7;

RA Mullered J., Hovav A.H., Nahary R., Fishan Y., Bercovier H.;

RT "Immunogenicity of a 16.7 kDa Mycobacterium paratuberculosis antigen.";

RL Microb. Pathog. 34:81-90(2003).

RN [2]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=K10;

RA Li L., Bannantine J., Zhang Q., Amosin A., Alt D., Kapur V.;

RL Submitted (SEP-2003) to the EMBL/Genbank/DBD databases.

DR EMBL; AY064405; AAL40871.1; -; Genomic_DNA.

DR EMBL; AB017233; AAS03970.1; -; Genomic_DNA.

DR HSSP; Q57549; 1Q98.

DR SMR; Q8GP25; 2-164.

DR GO; GO:0016491; P:oxidoreductase activity; IBA.

DR GO; GO:0009031; P:thiol peroxidase activity; IBA.

DR InterPro; IPR008866; Ahpc-TSA.

DR InterPro; IPR012336; Thioresdoxin-like.

DR InterPro; IPR012335; Thioresdoxin_fold.

DR InterPro; IPR002065; Tpx.

DR Pfam; PF00578; Ahpc-TSA; 1.

DR PROSITE; PS01265; Tpx; 1.

KM Complete proteome.

SQ SEQUENCE 164 AA; 16685 MW; 7F50AA6FEDC54545 CRC64;

Query Match 84.8%; Score 705; DB 2; Length 164;

Best Local Similarity 84.1%; Pred. No. 5,1e-53;

Matches 138; Conservative 12; Mismatches 14; Indels 0; Gaps 0

QY 1 MAQITLRGNAINTVGEELPAVGSAPAFITLTGGDLGVISDQFRGKSVLINTFPSTDPYC 60

Db 1 MAQITLRGNPINTVGEELPSVGSKAPAFSLTGGDLSPVTSQSGKPLVINTIPISTDPYC 60

QY 61 ATSVRTFDERAAASGATVLCYSKDLPPAOKRPGCAEGENTWPAFAFDSFSDYGVITIA 120

Db 61 ATSVRTFDERAAAGATVLANVSKDLPFAQAFPGCAEGENTWPAFAFDSFSDYGVITIA 120

QY 121 DGPMAGLARAIIVYGADGNVAYTELVETIAOEPYEAALALG 164

Db 121 DGPMAGLARAIIVVGADGNVAYTELVETIAOEPYEAALALAAAG 164

RESULT 5

084G72 MYCAV

ID 084G72 MYCAV PRELIMINARY; PRT; 160 AA.

AC 084G72;

DT 01-JUN-2003 (TRMBLrel. 24, Created)

DT 01-JUN-2003 (TRMBLrel. 24, Last sequence update)

DT 01-OCT-2003 (TRMBLrel. 25, Last annotation update)

DE Putative thiol peroxidase (Fragment).

OS Mycobacterium avium.

OC Bacteria; Actinobacteria; Actinomycetales;

OC Corynebacteriaceae; Mycobacteriaceae; Mycobacterium;

OC Mycobacterium avium complex (MAC).

OX NCBI_TaxID=1764;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RA Mullered J., Bercovier H.;

RL	Submitted (Oct-2002) to the EMBL/Genbank/DBJ databases.
DR	EMBL; AY16763; AAC01122.1; -, genomic_DNA.
DR	HSSP; Q57549; 1098.
DR	SME; Q84G72; 2-160.
DR	GO; GO:0009031; F:thiol peroxidase activity; IEA.
DR	InterPro; IPR000866; Ahpc-TSA.
DR	InterPro; IPR012336; Thioresdoxin-like.
DR	InterPro; IPR012335; Thioresdoxin_fold.
DR	InterPro; IPR02065; TPX.
DR	Pfam; PF00578; Ahpc-TSA; 1.
DR	PROSITE; PS01265; TPX; 1.
KW	Peroxidase.
FT	NON TER
SO	SEQUENCE
Query Match	160 AA; 160
Best Local Similarity	84.4%; Score 701; DB 2; Length 160;
Matches 137; Conservative	85.6%; Pred.No. 1.le-52;
	Matches 11; Mismatches 12; Indels 0; Gaps 0;
Dy	1 MAQITLKGNAINTVGELPAVGSPPAPATLTNGDLCVTSDDFRGKSVTLNTPFSVDTPVC 60
Dd	1 MAQITLKGNPNINTVGELPSVGSKAFAFSLTGGSDSPVDSFGSGPLVNTFPFSIDTPVC 60
Dy	61 ATSVTRDERAAASGATVLCVSXOLPFPOKRCFCGAEGEENMPASAFRDSGEEDGVYLA 120
Dd	61 ATSVRTFERRAAGATVAVLVNSKOLFPRQARFCGAEGLLENKYTKIASAFRDSFGDDGVYLA 120
Dy	121 DGPMAGLARAIVVIGADGNVAYTELVPEIAQBENYEAL 160
Dd	121 DGPMAGLLARAIVVVGADGNVAYTELVPEIAQBENYEAL 160
RESULT 6	
Q6AA71_PROAC	
ID	Q6AA71 PROAC PRELIMINARY; PRT; 167 AA.
AC	Q6AA71;
DT	25-OCT-2004 (TREMBLrel. 28. Created)
DT	25-OCT-2004 (TREMBLrel. 28. Last sequence update)
DT	25-OCT-2004 (TREMBLrel. 28. Last annotation update)
DB	Thiol peroxidase (EC 1.11.1.-).
GN	OrderedLocustNames=PPA0590;
OS	Propionibacterium acnes.
CC	Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC	Propionibacteriaceae; Propionibacteriaceae; Propionibacterium.
OX	NCBI_TaxID=1747;
RN	[1]
RP	NUCLEOTIDE SEQUENCE.
RC	STRAIN-KPA171202 / DSM 16379;
RX	PubMed=15286373; DOI=10.1126/science.1100330;
RA	Brueggemann H., Henne A., Hoester F., Liesegang H., Wlezer A.,
RA	Strittmatter A., Hugler S., Duerer J., Gottschalk G.;
RT	"The complete genome sequence of Propionibacterium acnes, a commensal
RT	of human skin."
RU	Science 305:671-673(2004).
DR	EMBL; AB017283; AAT82345.1; -, Genomic_DNA.
DR	GO; GO:0016491; F:oxygenoreductase activity; IEA.
DR	GO:0009031; F:thiol peroxidase activity; IEA.
DR	InterPro; IPR000866; Ahpc-TSA.
DR	InterPro; IPR012336; Thioresdoxin-like.
DR	InterPro; IPR02065; TPX.
DR	Pfam; PF00578; Ahpc-TSA; 1.
DR	PROSITE; PS01265; TPX; 1.
KW	Complete proteome; Oxidoreductase; Peroxidase.
SO	SEQUENCE
Query Match	62.5%; Score 519.5; DB 2; Length 167;
Best Local Similarity	63.8%; Pred. No. 5.8e-37;
Matches 104; Conservative	20; Mismatches 38; Indels 1; Gaps 1;
Dy	1 MAQITLKGNAINTVGELPAVGSPPAPATLTNGDLCVTSDDFRGKSVTLNTPFSVDTPVC 60
Dd	1 MATTAFFGKFLNVTGDLPGVGSLLPSTLVKSDLSRLSRDELKGRCLVNIFFPSVDTCVC 60

[illegible]

RESULT 7

ID	Q7NZQ1_CHRVO PRELIMINARY,	PRT;	165 AA.
AC	Q7NZQ1_		
DT	01-MAR-2004 (TrEMBLrel. 26, Created)		
DT	01-MAR-2004 (TrEMBLrel. 26, last sequence update)		
DT	01-MAR-2004 (TrEMBLrel. 26, last annotation update)		
DE	Thiol peroxidase (EC 1.11.1.-).		
GN	Name=tpx; OrderedAccessionNames=CV0870;		
OS	Chromobacterium violaceum.		
OC	Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;		
OC	Neisseriaceae; Chromobacterium.		
OX	NCBI_Taxid=536;		

RP NUCLEOTIDE SEQUENCE.
 RX STRAIN=ATCC 12472 / DSM 30191;
 RX MEDLINE=22882880; PubMed=14500782; DOI=10.1073/pnas.1832124100;
 RA Vaeconcelos A.T.R., de Almeida D.F., Hungria M., Guimaraes C.T.,
 RA Antonio R.V., Almeida F.C., de Almeida L.G.P., de Almeida R.,
 RA Alves-Gomes J.A., Andrade E.M., Arrêpe J., de Araujo M.F.F.,
 RA Ascoli-Filho S., Azevedo V., Baptista A.J., Batuas L.A.M.,
 RA Batista J.S., Belo A., van den Berg C., Bogo M., Bonatto S.,
 RA Bordignon J., Brigido M.M., Brito C.A., Brocchi M., Burity H.A.,
 RA Camargo A.A., Cardoso D.D.F., Carmiro N.P., Carraro D.M.,
 RA Carvalho C.M.B., Cascardo J.C.M., Cavada B.S., Chueire L.M.O.,
 RA Czezynek1-Pasa T.B., Cunha-Junior N.C., Fagundes N., Falcao C.L.,
 RA Fantiatti P., Farias I.P., Felipe M.S.S., Ferrari L.P., Ferro J.A.,
 RA Fetto M.I.T., Freitas N.S.A., Furian L.R.,
 RA Gazzinelli R.T., Gomes E.A., Goncalves P.R., Grangeiro T.B.,
 RA Grattapaglia D., Grlard E.C., Hanna E.S., Jardim S.N., Laurino J.,
 RA Leal L.C.T., Lima L.P.A., Loureiro M.P., Lyra M.C.C.P.,
 RA Madrita H.M.P., Mantlo G.P., Maranhao A.O., Martins W.S.,
 RA di Mauro S.M.Z., de Medeiros S.R.B., Melsenner R.V., Moreira M.A.M.,
 RA Nascimento P.F., Nicolas M.F., Oliveira J.G., Oliveira S.C.,
 RA Patkxo R.F.C., Parente J.A., Pedrosa F.O., Pena S.D.J., Pereira J.O.,
 RA Pereira M., Pinto L.S.R.C., Pinto L.S., Porto J.I.R., Potrich D.P.,
 RA Ramalho-Neco C.B., Reis A.M.M., Rigo L.U., Rondinelli E.,
 RA Santos E.B.P., Santos F.R., Schneider M.P.C., Senzano H.N.,
 RA Silva A.M.R., da Silva A.L.C., Silva D.W., Silva R., Simoes I.C.,
 RA Simon D., Soares C.M.A., Soares R.B.A., Souza E.M., Souza K.R.L.,
 RA Souza R.C., Steffens M.B.R., Steindl M., Teixeira S.R., Umenyi T.,
 RA Vettore A., Wassen R., Zaha A., Simpson A.J.G.;
 RT "the complete genome sequence of *Chromobacterium violaceum* reveals
 RT remarkable and exploitable bacterial adaptability.";
 RL Proc. Natl. Acad. Sci. U.S.A. 100:11660-11665(2003).
 DR EMBL; AE016912; AA058545.1; -; Genomic_DNA.
 DR HSSP; Q57549; 1038.
 DR GO; GO:0016491; F:oxidoreductase activity; IEA.
 DR GO; GO:0009031; F:thiol peroxidase activity; IEA.
 DR InterPro; IPR0000866; Anpc-TSA.
 DR InterPro; IPR012336; Thioledoxin-like.
 DR InterPro; IPR012335; Thioledoxin_fold.
 DR InterPro; IPR002065; TPX.
 DR Pfam; PF00578; Anpc-TSA; 1.
 DR PROSITE; PS01265; TPX; 1.
 KW Complete proteome; Oxidoreductase; Peroxidase.
 KW SEQUENCE 165 AA; 17002 MW; F01AB34A5A637206 CRC64;

Db	1	MAVVTLRGNPVDVAGLIPAKKGEOAPALQITGADLNAVSLASVAGRRKLIINPSPVDPTC	60
Oy	61	ATSVRTFDERRAAASGAT-VLCSKDLPEAKRRCFGAEETENVMPASAFRD-SFEEDGVT	118
Db	61	ATSVRTFDERRAAALADTVLVCSSADLPFGQRFCGAEGIENVVLTSTFRNGSPFAAVGVK	120
Oy	119	IADGPMAGLAARLVITGADGNVATELVPELADPNTFAALLAAL	163
Db	121	IASGPIASGLPARRVVVLDADVDYLSQLGSIADSPDTPAALLAAL	165

RESULT 8

ID	TPX	CHL/TE	STANDARD;	PRT;	168 AA.
AC	Q8KED5				
DT	28-FEB-2003	(Rel. 41, Created)			
DT	28-FEB-2003	(Rel. 41, Last sequence update)			
DT	13-SEP-2005	(Rel. 48, Last annotation update)			
DE	(Probable thiol peroxidase (EC 1.11.1.-)).				
GN	Name=Lpx;	OrderedLocustNames=CT0754;			
OS	Chlorobium	tepidum.			
OC	Bacteria;	Chlorobi;	Chlorobia;	Chlorobiales;	Chlorobiaceae
CC					Chlorobaculum.
CK	NCBI_TaxId=1097;				

```

RC NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RP STAIN-TLS / ATCC 49652 / DSM 102025;
RX MEDLINE=22103685; PubMed=12093901; DOI=10.1073/pnas.132181499;
RA Eisen J.A., Nelson K.E., Paulsen I.T., Heidelberg J.P., Wu M.,
RA Dodson R.J., Deboy K.T., Gwinn M.L., Nelson W.C., Haft D.H.,
RA Hickey E.K., Peterson J.D., Durkin A.S., Kolonay J.F., Yang F.,
RA Holt I.E., Umayam L.A., Mason T.M., Brenner M., Shea T.P.,
RA Parksey D.S., Niernan W.C., Feldlyum T.V., Hansen C.L., Craven M.B.,
RA Radune D., Vamathevan J.J., Khouri H.M., White O., Gruber T.M.,
RA Kerchun K.A., Venter J.C., Tettelin H., Bryant D.A., Fraser C.M.,
RT "the complete genome sequence of Chlorobium tepidum TLS, a
RT photosynthetic, anaerobic, green-sulfur bacterium." ;
RL Proc. Natl. Acad. Sci. U.S.A. 99:9509-9514(2002) .
CC -!- FUNCTION: Has antioxidant activity. Could remove peroxides or
CC H2O(12) (By similarity).
CC -!- SIMILARITY: Belongs to the ahpc/TSA family. Tpx subfamily.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outbreaken -
CC the European Bioinformatics institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
DR EMBL; AE006470; AAAT1991.1; -; genomic_DNA.
DR HSSP; P72500; 1PSQ.
DR TIGR; CT0754; -.
DR HAMAP; MF_00269; -: 1.
DR InterPro; IPRO000866; Ahpc-TSA.
DR InterPro; IPRO12336; thioredoxin-like.
DR InterPro; IPRO12335; thioredoxin_fold.
DR InterPro; IPRO02065; tpx.
DR PANTHER; PTHR10681:SFA; tpx; 1.
DR Pfam; PF00578; Ahpc-TSA; 1.
DR PROSITE; PS01265; TPX; 1.
KM Antioxidant; Complete proteome; Oxidoreductase; Peroxidase.
SQ SEQUENCE 168 AA; 17714 MW; 4B84D04B5028073 CRC64;

Query Match          60.0%; Score 499; DB 1; Length 168;
Best Local Similarity 60.6%; Pred. No. 3,4e-35;
Matches 100; Conservative 20; Mismatches 43; Indels 2; Gaps 2;

Oy      1 MAQTLLGNAINTVGEIIPAVGSPAPAFITLGGLGVTSDDPFKSGVLNIFSVDTPYVC 60
        |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      3 MATTLTKNGSIHTIGEELPAVGSQLPATFLTVKSIDSSEVS PADFPACKCVLNTPSELDAVC 62

Oy      61 ATSTRTDERAASAG-ATVLCVSDLPFPAORCFGCAGSTENWVASAFRD-SPEEDYGYT 118
        |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      63 AASYRRFFKEKGGERGDAAVLTISDLDPFAQCRFTTEGLDNVVSLAYTRSPERGLDGLT 122

```


Qy 119 IADGPMAGLARAIVTIGADGNVATYELVPEIAOBPNYEAALAL 163
 Db 123 ITDGPLKGLSLRAVIVTADSGKVLVAFQVPRIVQPDVDAALAL 167

RESULT 9

TPX_PSEAE STANDARD; PRT; 165 AA.
 AC P57668;
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 10-MAY-2005 (Rel. 47, Last annotation update)
 DE Probable thiol peroxidase (EC 1.11.1.-).
 OS Name=tpx; OrderedLocustNames=PA2532;
 OS Pseudomonas aeruginosa.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 OC Pseudomonadaceae; Pseudomonas.
 NCBI_TaxID=287;
 RN [1]
 RP NCBIOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=ATCC 15692 / PAOI;
 RX MEDLINE=20437337; PubMed=10984043; DOI=10.1038/35023079;
 RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warrenner P.,
 RA Hickey M.J., Brinkman F.S.L., Hutnagle W.O., Kowalik D.J., Lagrou M.,
 RA Garber R.L., Goltz L., Tolentino E., Westbrock-Wadman S., Yuan Y.,
 RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Landig K., Lim R.M.,
 RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
 RA Reizer J., Slier M.H.Jr., Hancock R.E.W., Lory S., Olson M.V.,
 RT "Complete genome sequence of Pseudomonas aeruginosa PAOI, an
 RT opportunistic pathogen.";
 RL Nature 406:959-964(2000).
 CC -1- FUNCTION: Has antioxidant activity. Could remove peroxides or
 CC H(2)O(2) (by similarity).
 CC -1- SIMILARITY: Belongs to the ahpc/TSA family. Tpx subfamily.

CC -----
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.

CC
 DR EMBL; AB004681; AAG05920.1; -; Genomic_DNA.
 DR PIR; E83328; E83328.
 DR HSSP; P72500; 1PSQ.
 DR HAMAP; MP_00269; 1.
 DR InterPro; IPR000866; Ahpc-TSA.
 DR InterPro; IPR012336; Thioresdoxin-like.
 DR InterPro; IPR012335; Thioresdoxin-like.
 DR InterPro; IPR002065; Tpx.
 DR PANTHER; PTHR10681.SF4; Tpx; 1.
 DR Pfam; PF00578; Ahpc-TSA; 1.
 DR PROSITE; PS01265; Tpx; 1.
 KW Antioxidant; Complete proteome; Oxidoreductase; Peroxidase.
 SQ SEQUENCE 165 AA; 17234 MM; 6CBG3A350B57A478 CRC64;

Query Match 58.5%; Score 486; DB 1; Length 165;
 Best Local Similarity 60.0%; Pred. No. 4.5e-34;
 Matches 99; Conservative 20; Mismatches 44; Indels 2; Gaps 2;

Qy 1 MAQITLGNAINVYGLPAVSPAPAFITLNGDVGIVISDQFRGKSVLNIIPSVDPVC 60
 Db 1 MAQITLGNAINVYGLPAVSPAPAFITLNGDVGIVISDQFRGKSVLNIIPSVDPVC 60
 Qy 61 ATSVRTED-ERAAASGATVLCVSKDLPFAQKRFPGABGTENVMPASAFRD-SFGEDYGV 118
 Db 61 ATSVRKKNVBAKGLANTVLCISADLPFAQKRFPGABGTENVMPASAFRD-SFGEDYGV 120
 Qy 119 IADGPMAGLARAIVTIGADGNVATYELVPEIAOBPNYEAALAL 163
 Db 121 IASGPLAGLARAIVVLDQNKVLSHLSVGEIADEPNYEAALAL 165

RESULT 10
 Q7MU37_PORGI PRELIMINARY; PRT; 179 AA.
 AC Q7MU37;
 DT 01-MAR-2004 (Tremblrel. 26, Created)
 DT 01-MAR-2004 (Tremblrel. 26, Last sequence update)
 DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
 DE Thiol peroxidase.
 OS OrderedLocustNames=Pg1729;
 OS Porphyromonas gingivalis (Bacteroides gingivalis).
 OC Bacteria; Bacteroidetes; Bacteroidetes (class); Bacteroidales;
 OC Porphyromonadaceae; Porphyromonas.
 NCBI_TaxID=837;
 RN [1]
 RP NCBIOTIDE SEQUENCE.
 RC STRAIN=M83;
 RX MEDLINE=22829867; PubMed=12949112;
 RX DOI=10.1128/JB.185.18.5591-5601.2003;
 RA Nelson K.E., Fleischmann R.D., DeBoy R.T., Paulsen I.T., Fouts D.E.,
 RA Eisen J.A., Daugherty S.C., Dodson R.J., Durkin A.S., Gwinn M.L.,
 RA Haft D.H., Kolonay J.F., Nelson W.C., Mason T.M., Tallon L., Gray J.,
 RA Granger D., Tettelin H., Dong H., Galvan J.L., Duncan M.J.,
 RA Dewhirst F.E., Fraser C.M.;
 RT "Complete genome sequence of the oral pathogenic bacterium
 RT Porphyromonas gingivalis strain W83.";
 RL J. Bacteriol. 185:5591-5601(2003).
 DR EMBL; AB017178; AAO6736.1; -; Genomic_DNA.
 DR HSSP; P37901; 10XH.
 DR TIGR; PG1729; -.
 DR GO; GO:0016491; P:oxidoreductase activity; IEA.
 DR GO; GO:0009031; P:thiol peroxidase activity; IEA.
 DR InterPro; IPR000866; Ahpc-TSA.
 DR InterPro; IPR012336; Thioresdoxin-like.
 DR InterPro; IPR012335; Thioresdoxin-like.
 DR InterPro; IPR002065; Tpx.
 DR Pfam; PF00578; Ahpc-TSA; 1.
 DR PROSITE; PS01265; Tpx; 1.
 KW Complete proteome; Peroxidase.
 SQ SEQUENCE 179 AA; 19188 MM; 1C68D2F6C0652CC1 CRC64;

Query Match 57.5%; Score 477.5; DB 2; Length 179;
 Best Local Similarity 57.8%; Pred. No. 2.7e-33;
 Matches 96; Conservative 25; Mismatches 42; Indels 3; Gaps 3;

Qy 1 MAQITLGNAINVYGLPAVSPAPAFITLNGDVGIVISDQFRGKSVLNIIPSVDPVC 59
 Db 13 MATVTLGKINININGQLPVGSAVAPDFKAVRADLSVLSSEFGKAVINIPFSIDTV 72
 Qy 60 CATSVRTEDERAAA-SGATVLCVSKDLPFAQKRFPGABGTENVMPASAFRD-SFGEDYGV 117
 Db 73 CAASVRRFNGEASLDNTVVLCISKDLPFAQKRFPGABGTENVMPASAFRD-SFGEDYGV 132
 Qy 118 TIADGPMAGLARAIVTIGADGNVATYELVPEIAOBPNYEAALAL 163
 Db 133 LMTDGPLKGLARAIVVLDQNKVLSHLSVGEIADEPNYEAALAL 178

RESULT 11

Q4KCR9_PSEBS PRELIMINARY; PRT; 166 AA.
 AC Q4KCR9;
 DT 13-SEP-2005 (Tremblrel. 31, Created)
 DT 13-SEP-2005 (Tremblrel. 31, Last sequence update)
 DT 13-SEP-2005 (Tremblrel. 31, Last annotation update)
 DE Thiol peroxidase (EC 1.11.1.-).
 OS ORFNames=pfL_2858;
 OS Pseudomonas fluorescens (strain Pf-5).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 OC Pseudomonadaceae; Pseudomonas.
 NCBI_TaxID=220664;
 RN [1]
 RP NCBIOTIDE SEQUENCE.
 RC STRAIN=PF-5;

Qy 1 MAQITLGNAINVYGLPAVSPAPAFITLNGDVGIVISDQFRGKSVLNIIPSVDPVC 60
 Db 1 MAQITLGNAINVYGLPAVSPAPAFITLNGDVGIVISDQFRGKSVLNIIPSVDPVC 60
 Qy 61 ATSVRTED-ERAAASGATVLCVSKDLPFAQKRFPGABGTENVMPASAFRD-SFGEDYGV 118
 Db 61 ATSVRKKNVBAKGLANTVLCISADLPFAQKRFPGABGTENVMPASAFRD-SFGEDYGV 120
 Qy 119 IADGPMAGLARAIVTIGADGNVATYELVPEIAOBPNYEAALAL 163
 Db 121 IASGPLAGLARAIVVLDQNKVLSHLSVGEIADEPNYEAALAL 165

RK PubMed=15980861; DOI=10.1038/nbt1110;
 RA Paulsen I.T., Press C., Ravel J., Kobayashi D., Myers G.S.,
 RA Mavrodidi D., DeBoy R.T., Seshadri R., Ren Q., Madupu R., Dodson R.J.,
 RA Durkin S., Brinac L.M., Daugherty S.C., Sullivan S.A., Rosovitz M.,
 RA Gelin M.L., Zhou L., Nelson W.C., Weidman J., Watkins K., Tran K.,
 RA Khouri H.M., Pierson E., Pierson L. III, Thomashow L., Loper J.,
 RT "Complete genome sequence of the plant commensal *Pseudomonas*
 RT *fluorescens* Pf-5."
 RL EMBL; BioTechnol. 23:873-878(2005).
 DR EMBL; CP000076; AA92130.1; -; Genomic_DNA.
 KW Oxidoreductase; Peroxidase.
 SQ SEQUENCE 166 AA; 17418 MW; 198EC133AA21B56 CRC64;

Query Match 57.3%; Score 476; DB 2; Length 166;
 Best Local Similarity 59.4%; Pred. No. 3.3e-33;
 Matches 98; Conservative 19; Mismatches 46; Indels 2; Gaps 2;

QY 1 MAQITLRGNAINVGEIPAVGSPAPATLTGGLGVYSSDQFRGKSVLTNIFFSVDTPTVC 60
 DB 1 MAQITLRGNAINVGEIPAVGSPAPATLTGGLGVYSSDQFRGKSVLTNIFFSVDTPTVC 60
 QY 61 ATSVRTFDEBAA-AASGATV.CVS.KDLPFAQKRCGAGSTENVMASAFRD-SFGEDYGT 118
 DB 61 ATSVRKFNMAQANLANTVLCISADLPFAQKRCGAGSTENVMASAFRD-SFGEDYGT 120

QY 119 ADGPMAGLLARAIVIGADGNVATYELVPEIAQEPNYEALAL 163
 DB 121 ADGPMAGLLARAIVIGADGNVATYELVPEIAQEPNYEALAL 165

RESULT 12
 TPX CORGL
 ID TPX CORGL STANDARD; PRT; 165 AA.

AC OENRG3;
 DT 28-FEB-2003 (Rel. 41, Last Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 10-MAY-2005 (Rel. 47, Last annotation update)
 DN Probable thiol peroxidase (EC 1.11.1.-).
 GN Name:tpx; OrderedLocNames=Cg11086, Cg1236;
 OS *Corynebacterium glutamicum* (Brevibacterium *fluvum*).
 OC Bacteria; Actinobacteria; Actinobacteriales; Actinomycetales;
 OC *Corynebacteriaceae*; *Corynebacterium*.
 OC *Corynebacterium*.
 OX NCBI_TaxID=1718;
 RN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RP STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;
 RC Nakagawa S.;
 RA "Complete genomic sequence of *Corynebacterium glutamicum* ATCC 13032.",
 RT Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
 RL [2]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;
 RX PubMed=12948626; DOI=10.1016/S0168-1656(03)00154-8;
 RA Kallinowski J., Bathe B., Bartels D., Bischoff N., Bott M.,
 RA Burkovski A., Dusch N., Eggeling L., Kikmanns B.J., Galsalat L.,
 RA Goessmann A., Hartmann M., Huthmacher K., Kraemer R., Linke B.,
 RA Mohardy A.C., Meyer F., Moeckel B., Pfeifferle W., Pfeiler A.,
 RA Rey D.A., Rueckert C., Rupp O., Sahn H., Wendisch V.F., Wiegand I.,
 RA Tauch A.;
 RT "The complete *Corynebacterium glutamicum* ATCC 13032 genome sequence
 RT and its impact on the production of L-aspartate-derived amino acids
 RT and vitamins".
 RL J. Biotechnol. 104:5-25(2003).
 CC -1- FUNCTION: Has antioxidant activity. Could remove peroxides or
 CC H(2)O(2) (by similarity).
 CC -1- SIMILARITY: Belongs to the ahpc/TSA family. Tpx subfamily.

CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.

DR EMBL; BA000036; BAB98479.1; -; Genomic_DNA.
 DR EMBL; BX927151; CAF19791.1; ALT_INIT; Genomic_DNA.
 DR HSPB; P72500; 1PSQ.
 DR HAMAP; MF 00269; -; 1.
 DR InterPro; IPR000866; Ahpc-TSA.
 DR InterPro; IPR012335; Thioresoxin-1like.
 DR InterPro; IPR012335; Thioresoxin-1like.
 DR InterPro; IPR002065; Tpx.
 DR PANTHER; PTHR10681.SF4; Tpx; 1.
 DR Pfam; PF00578; Ahpc-TSA; 1.
 DR PROSITE; PS01265; Tpx; 1.
 KW Antioxidant; Complete proteome; Oxidoreductase; Peroxidase.
 SQ SEQUENCE 165 AA; 17532 MW; 8A13B971070B2CDB CRC64;

Query Match 57.0%; Score 473.5; DB 1; Length 165;
 Best Local Similarity 57.3%; Pred. No. 5.4e-33;
 Matches 94; Conservative 24; Mismatches 45; Indels 1; Gaps 1;

QY 1 MAQITLRGNAINVGEIPAVGSPAPATLTGGLGVYSSDQFRGKSVLTNIFFSVDTPTVC 60
 DB 1 MAQITLRGNAINVGEIPAVGSPAPATLTGGLGVYSSDQFRGKSVLTNIFFSVDTPTVC 60
 QY 61 ATSVRTFDEBAA-AASGATV.CVS.KDLPFAQKRCGAGSTENVMASAFRD-SFGEDYGT 119
 DB 61 ATSVRKFNMAQANLANTVLCISADLPFAQKRCGAGSTENVMASAFRD-SFGEDYGT 120

QY 120 ADGPMAGLLARAIVIGADGNVATYELVPEIAQEPNYEALAL 163
 DB 121 ADGPMAGLLARAIVIGADGNVATYELVPEIAQEPNYEALAL 164

RESULT 13
 OSLBP6_BACFN
 ID OSLBP6_BACFN PRELIMINARY; PRT; 166 AA.
 AC OSLBP6;
 DT 01-FEB-2005 (TrEMBLrel. 29, Created)
 DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
 DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
 DN Putative thiol peroxidase (EC 1.11.1.-).
 GN Name:tpx; OrderedLocNames=BF2777;
 OS *Bacteroides fragilis* (strain ATCC 25285 / NCTC 9243).
 OC Bacteria; Bacteroidetes; Bacteroidetes (class); Bacteroidales;
 OC *Bacteroides*; *Bacteroides*.
 OX NCBI_TaxID=272559;
 RN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RP Cordenno-Tarraga A.-M., Patrick S., Crossman L.C., Blakely G.,
 RX PubMed=15746427; DOI=10.1126/science.1107008;
 RA Abrate V., Leonard N., Foxton I., Duerden B., Harris B., Quail M.A.,
 RA Barton A., Clark L., Corton C., Doggett J., Holden M.T.G., Larke N.,
 RA Lane A., Lord A., Norbertczak H., Ormond D., Price C.,
 RA Rabinowitsch E., Woodward J., Barrell B.G., Partridge J.;
 RT "Extensive DNA inversions in the *B. fragilis* genome control variable
 RT gene expression".
 RL Science 307:1463-1465(2005).
 DR EMBL; CR626927; CAH08472.1; -; Genomic_DNA.
 DR InterPro; IPR000866; Ahpc-TSA.
 DR InterPro; IPR012336; Thioresoxin-1like.
 DR InterPro; IPR012335; Thioresoxin-1like.
 DR InterPro; IPR002065; Tpx.
 DR Pfam; PF00578; Ahpc-TSA; 1.
 DR PROSITE; PS01265; Tpx; 1.
 KW Complete proteome; Oxidoreductase; Peroxidase.
 SQ SEQUENCE 166 AA; 17938 MW; DE898C5407EAB68 CRC64;

Query Match 56.2%; Score 467; DB 2; Length 166;
 Best Local Similarity 59.4%; Pred. No. 2e-32;
 Matches 98; Conservative 18; Mismatches 47; Indels 2; Gaps 2;

QY 1 MAQITLRGNAINVGEIPAVGSPAPATLTGGLGVYSSDQFRGKSVLTNIFFSVDTPTVC 60
 DB 1 MAQITLRGNAINVGEIPAVGSPAPATLTGGLGVYSSDQFRGKSVLTNIFFSVDTPTVC 60

QY 61 ATSVRTFDERAASGAT-VLCVSKDLPFAQKRFCAEGSTENVMPASAFRDS-FGEDYGVY 118
 DB 61 ATSVRKFKMAAGKMDYVLAISKDLPFAQGRFCTTBGIENTVPLSDRFRFSDPDESIGVR 120
 QY 119 IADGPMAGLLARAVIVGADGNVATYELVPEPIAQPENYKALAL 163
 DB 121 MADGPPLAGLLARAVVIGKDGKVAITELVPEITQEPDYEKALAAV 165

RESULT 14

064SL6_BACFR
 ID 064SL6_BACFR PRELIMINARY; PRT; 166 AA.
 AC 064SL6;
 DT 25-OCT-2004 (TREMBlrel. 28, Created)
 DT 25-OCT-2004 (TREMBlrel. 28, Last sequence update)
 DT 25-OCT-2004 (TREMBlrel. 28, Last annotation update)
 DE Putative thiol peroxidase.
 GN OrderedLocustNames=BP2763;
 OS Bacteroides fragilis.
 OC Bacteroidetes; Bacteroidetes (class); Bacteroidales;
 OC Bacteroidaceae; Bacteroides.
 OX NCBI_TaxID=817;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=YCH46;
 RX PubMed=15466707; DOI=10.1073/pnas.0404172101;
 RA Kuwahara T., Yamashita A., Hiraoka H., Nakayama H., Toh H., Okada N.,
 RA Kohara S., Hattori M., Hayashi T., Ohnishi Y.,
 RT Genomic analysis of Bacteroides fragilis reveals extensive DNA
 RT Inversions regulating cell surface adaptation."
 RL Proc. Natl. Acad. Sci. U.S.A. 101:14919-14924(2004).
 DR EMBL; AP006841; BAD49513.1; -? Genomic DNA.
 DR GO; GO:0016491; F:oxidoreductase activity; IEA.
 DR InterPro; IPR009031; F:thiol peroxidase activity; IEA.
 DR InterPro; IPR00866; AhpC-TSA.
 DR InterPro; IPR012336; Thioresoxin-like.
 DR InterPro; IPR02065; TPX.
 DR Pfam; PF00578; AhpC-TSA; 1.
 DR PROSITE; PS01265; TPX; 1.
 KM Complete proteome; Peroxidase.
 SQ SEQUENCE 166 AA; 17938 MW; D898C54074B868 CRC64;

Query Match 56.2%; Score 467; DB 2; Length 166;
 Best Local Similarity 59.4%; Pred. No. 2e-32;
 Matches 98; Conservative 18; Mismatches 47; Indels 2; Gaps 2;

QY 1 MAQITLGNAINTVGELPAVGSPPAPFTLTCGDLGVISDQFRGKSVLNTFFPSVDTPVC 60
 DB 1 MATTNFPGQPVKLTIGEFIOVGKVAPELVKSDLSFPAKDKKNTVLTNIFPSLDTGVC 60
 QY 61 ATSVRTFDERAASGAT-VLCVSKDLPFAQKRFCAEGSTENVMPASAFRDS-FGEDYGVY 118
 DB 61 ATSVRKFKMAAGKMDYVLAISKDLPFAQGRFCTTBGIENTVPLSDRFRFSDPDESIGVR 120
 QY 119 IADGPMAGLLARAVIVGADGNVATYELVPEPIAQPENYKALAL 163
 DB 121 MADGPPLAGLLARAVVIGKDGKVAITELVPEITQEPDYEKALAAV 165

RESULT 15

06AM98_DSPPS
 ID 06AM98_DSPPS PRELIMINARY; PRT; 167 AA.
 AC 06AM98;
 DT 25-OCT-2004 (TREMBlrel. 28, Created)
 DT 25-OCT-2004 (TREMBlrel. 28, Last sequence update)
 DT 25-OCT-2004 (TREMBlrel. 28, Last annotation update)
 DE Probable thiol peroxidase.
 GN OrderedLocustNames=DP1798;
 OS Desulfotalea psychrophila.
 OC Bacteria; Proteobacteria; Deltaproteobacteria; Desulfobacteriales;
 OC Desulfobacteraceae; Desulfotalea.
 OX NCBI_TaxID=84980;
 RN [1]

RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=LSv54 / DSM 12343;
 RX PubMed=15305914; DOI=10.1111/j.1462-2920.2004.00665.x;
 RA Babus R., Ruepp A., Frickey T., Rattei T., Fattmann B., Stark M.,
 RA Bauer M., Zibat A., Lombardoc T., Becker I., Amann U., Gellner K.,
 RA Teeling H., Leuschner W.D., Gloeckner F.-O., Lupas A.N., Amann R.,
 RA Klenk H.-P.;
 RT "The genome of Desulfotalea psychrophila, a sulfate-reducing bacterium
 RT from permanently cold Arctic sediments."
 RL Environ. Microbiol. 6:887-902(2004).
 DR EMBL; CR522870; CAG36527.1; -? Genomic DNA.
 DR GO; GO:0016491; F:oxidoreductase activity; IEA.
 DR GO; GO:0009031; F:thiol peroxidase activity; IEA.
 DR InterPro; IPR00866; AhpC-TSA.
 DR InterPro; IPR012336; Thioresoxin-like.
 DR InterPro; IPR02065; TPX.
 DR Pfam; PF00578; AhpC-TSA; 1.
 DR PROSITE; PS01265; TPX; 1.
 KM Complete proteome.
 SQ SEQUENCE 167 AA; 17796 MW; 20A62BBF04BBD462 CRC64;

Query Match 56.0%; Score 465; DB 2; Length 167;
 Best Local Similarity 56.4%; Pred. No. 3e-32;
 Matches 93; Conservative 20; Mismatches 50; Indels 2; Gaps 2;

QY 1 MAQITLGNAINTVGELPAVGSPPAPFTLTCGDLGVISDQFRGKSVLNTFFPSVDTPVC 60
 DB 1 MTQVSLNGTITETIGDLPYQDMAADFSLTASDLSSTKTLDDYTGNTVLTNIFPSIDTPVC 60
 QY 61 ATSVRTFDERAASGAT-VLCVSKDLPFAQKRFCAEGSTENVMPASAFRDS-SFGEDYGVY 118
 DB 61 AMSVRKFNPSDAAAGKNTKVLCTISADLPFAHARFCDEGLKNVITPLSAFRSPNFGKDYGOT 120
 QY 119 IADGPMAGLLARAVIVGADGNVATYELVPEPIAQPENYKALAL 163
 DB 121 ITTGPLTGLARAVVITSPAGKILTYQVPEITQEPNYSKALLLL 165

Search completed: April 14, 2006, 17:32:05
 Job time : 62.8796 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocelebration Ltd.

OM protein - protein search, using sw model

Run on: April 14, 2006, 17:32:59 ; Search time 15.2571 Seconds
(without alignments)
894.107 Million cell updates/sec

Title: US-10-620-246-8

Perfect score: 831
Sequence: 1 MAQITLRGNAINTVGBLPAV.....LVPBIAQSPNYEALALGA 165

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA.*
1: /cgn2_6/ptodata/1/1aa/5_COMB.pep.*
2: /cgn2_6/ptodata/1/1aa/6_COMB.pep.*
3: /cgn2_6/ptodata/1/1aa/H_COMB.pep.*
4: /cgn2_6/ptodata/1/1aa/PTUS_COMB.pep.*
5: /cgn2_6/ptodata/1/1aa/RR_COMB.pep.*
6: /cgn2_6/ptodata/1/1aa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	831	100.0	165	US-09-050-739-8	Sequence 8, Appli
2	486	58.5	194	US-09-252-991A-31239	Sequence 31239, A
3	466.5	56.1	165	US-09-603-208A-64	Sequence 64, Appl
4	466.5	56.1	165	US-09-602-777A-172	Sequence 172, App
5	441	53.1	172	US-09-543-681A-4608	Sequence 4608, App
6	418.5	50.4	176	US-09-489-039A-7716	Sequence 7716, Ap
7	341.5	41.1	123	US-09-602-777A-174	Sequence 174, App
8	319.5	38.4	172	US-09-583-110-5217	Sequence 3217, Ap
9	319.5	38.4	201	US-09-107-433-5126	Sequence 5126, Ap
10	314	37.8	173	US-08-956-171R-5218	Sequence 5218, Ap
11	314	37.8	173	US-08-781-986A-5218	Sequence 5218, Ap
12	307	36.9	164	US-09-710-379-252	Sequence 252, App
13	307	36.9	168	US-09-134-001C-3764	Sequence 3764, App
14	302	36.3	171	US-09-902-540-15283	Sequence 15283, A
15	282.5	34.0	144	US-08-961-083-44	Sequence 44, Appl
16	282.5	34.0	144	US-09-536-784-44	Sequence 44, Appl
17	282.5	34.0	144	US-09-765-271-44	Sequence 44, Appl
18	282.5	34.0	144	US-09-765-272A-44	Sequence 44, Appl
19	195.5	23.5	108	US-08-936-165A-443	Sequence 443, App
20	181.5	21.8	110	US-09-134-000C-5730	Sequence 5730, App
21	154.5	18.6	274	US-09-902-540-16718	Sequence 16718, A
22	121.5	14.6	257	US-08-467-265-16	Sequence 16, Appl
23	121.5	14.6	257	US-08-467-265-16	Sequence 16, Appl
24	121.5	14.6	257	US-09-407-891-16	Sequence 16, Appl
25	121.5	14.6	257	US-09-375-907-5	Sequence 5, Appli
26	117.5	14.1	198	US-09-202-329-11	Sequence 11, Appli
27	117.5	14.1	199	US-08-602-262-2	Sequence 2, Appli

28	117.5	14.1	199	2	US-09-004-716-2	Sequence 2, Appli
29	117.5	14.1	199	2	US-09-477-510-2	Sequence 2, Appli
30	117.5	14.1	256	2	US-09-538-092-1039	Sequence 1039, App
31	117.5	14.1	256	2	US-09-949-016-6760	Sequence 6760, App
32	117.5	14.1	258	2	US-09-949-016-10249	Sequence 10249, A
33	114.5	13.8	198	2	US-09-538-092-1055	Sequence 1055, Ap
34	114.5	13.8	206	2	US-09-949-016-8377	Sequence 8377, Ap
35	114	13.7	196	2	US-09-538-092-593	Sequence 593, App
36	113.5	13.7	196	2	US-09-556-877-294	Sequence 294, App
37	113.5	13.7	196	2	US-09-620-412C-294	Sequence 294, App
38	113.5	13.7	196	2	US-09-598-419-294	Sequence 294, App
39	111.5	13.4	198	2	US-08-467-265-15	Sequence 15, Appl
40	111.5	13.4	198	2	US-08-467-265-15	Sequence 15, Appl
41	111.5	13.4	198	2	US-09-407-891-15	Sequence 15, Appl
42	111	13.4	147	2	US-09-513-999C-5696	Sequence 5696, Ap
43	111	13.4	199	1	US-08-299-162A-2	Sequence 2, Appli
44	111	13.4	199	1	US-08-467-265-14	Sequence 14, Appli
45	111	13.4	199	1	US-08-467-265-17	Sequence 17, Appl

ALIGNMENTS

RESULT 1						
US-09-050-739-8						
Sequence 8, Application US/09050739						
Patent No. 6641814						
GENERAL INFORMATION:						
APPLICANT: NIELSEN, Peter						
APPLICANT: NIEMSEN, Nikke						
APPLICANT: OETTINGER, Thomas						
APPLICANT: RASMUSSEN, Peter Blak						
APPLICANT: ROSENKRANDS, Ida						
APPLICANT: WELDLING, Karin						
APPLICANT: FLORIO, Walter						
TITLE OF INVENTION: DERIVED ACIDS FRAGMENTS AND POLYPEPTIDE FRAGMENTS						
TITLE OF INVENTION: DERIVED FROM M. TUBERCULOSIS						
FILE REFERENCE: 670001-2002.1						
CURRENT APPLICATION NUMBER: US/09/050,739						
CURRENT FILING DATE: 1998-03-30						
EARLIER APPLICATION NUMBER: 0376/97						
EARLIER FILING DATE: 1997-04-02						
EARLIER APPLICATION NUMBER: 1277/97						
EARLIER FILING DATE: 1997-11-10						
EARLIER APPLICATION NUMBER: 60/044,624						
EARLIER FILING DATE: 1997-04-18						
EARLIER APPLICATION NUMBER: 60/070,488						
EARLIER FILING DATE: 1998-01-05						
NUMBER OF SEQ ID NOS: 173						
SOFTWARE: Patentin Ver. 2.0						
SEQ ID NO 8						
LENGTH: 165						
TYPE: PRT						
ORGANISM: Mycobacterium tuberculosis						
US-09-050-739-8						
Query Match 100.0%; Score 831; DB 2; Length 165;						
Best Local Similarity 100.0%; Pred. No. 1.2e-85;						
Matches 165; Conservative 0; Mismatches 0; Indels 0; Gaps 0;						
Qy	1	MAQITLRGNAINTVGBLPAV	ATTTGGDLGVISDPFRGKSVLTNI	PPSVDTPVC	60	
Db	1	MAQITLRGNAINTVGBLPAV	SPAPATTLTGGDLGVISDPFRGKSVLTNI	PPSVDTPVC	60	
Qy	61	ATSVRTFDEAAASGATVLC	VSVDLPFAQRFCAEGSTENVM	PASARDSFGEDYGTIA	120	
Db	61	ATSVRTFDEAAASGATVLC	VSVDLPFAQRFCAEGSTENVM	PASARDSFGEDYGTIA	120	
Qy	121	DGPAGGLARAIIVTIGADG	VNATTELYPEIAQSPNYEAL	ALGA 165		
Db	121	DGPAGGLARAIIVTIGADG	VNATTELYPEIAQSPNYEAL	ALGA 165		

```
RESULT 2
US-09-252-991A-31239
; Sequence 31239, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 31239
; LENGTH: 194
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-31239

Query Match          58.5%; Score 486; DB 2; Length 194;
Best Local Similarity 60.0%; Pred. No. 1,3e-46;
Matches 99; Conservative 20; Mismatches 44; Indels 2; Gaps 2;

QY 1 MAQITRGNAINTVGEIPAVGSPAPFTLTGDLGVISDQFRGKSVLTNIPPSVDTTPVC 60
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
DB 30 MAQVTLKGNVYVDGQIPKGAQAPAFSLVGGDLADVTLENFAGKRKRVLTNIPPSVDTTPC 89
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
QY 61 ATSVRTFD-ERAAASGATVLCVSKDLFPQAKRFGCAEGTENWMPASAFRD-SFGEDYGV 118
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
DB 90 ATSVRKFNBAASLENTVLCISKDLFPALGRFCAGLEINVTVPASAFSTFGEDNGIVL 149
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
QY 119 IADGPAGLLARAIIVTGADGNVAYTELVPBIAQBPYEAALAL 163
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
DB 150 IASGPLAGLLARAIIVVDGQNKVLAHSELVIADEPNYAAALAL 194
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||

RESULT 3
US-09-603-208A-64
; Sequence 64, Application US/09603208A
; Patent No. 6822084
; GENERAL INFORMATION:
; APPLICANT: Pompeius, Markus
; APPLICANT: Kroeger, Burkhard
; APPLICANT: Schroeder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Heberhauser, Gregor
; APPLICANT: Lee, Hyung-Shick
; APPLICANT: Kim, Hyung-Joon
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING STRESS,
; TITLE OF INVENTION: RESISTANCE AND TOLERANCE PROTEINS
; FILE REFERENCE: BGI-124CP
; CURRENT APPLICATION NUMBER: US/09/603,208A
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: 60/141031
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: 60/142692
; PRIOR FILING DATE: 1999-07-01
; PRIOR APPLICATION NUMBER: 60/151214
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: DE 19930429.7
; PRIOR FILING DATE: 1999-07-01
; PRIOR APPLICATION NUMBER: DE 19931413.6
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931457.8
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931541.8
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19932209.0
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932230.9
```

```
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932914.1
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19940764.9
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: DE 19941382.7
; PRIOR FILING DATE: 1999-08-31
; NUMBER OF SEQ ID NOS: 306
; SEQ ID NO 64
; LENGTH: 165
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-603-208A-64

Query Match          56.1%; Score 466.5; DB 2; Length 165;
Best Local Similarity 56.7%; Pred. No. 1.4e-44;
Matches 93; Conservative 23; Mismatches 47; Indels 1; Gaps 1;

QY 1 MAQITRGNAINTVGEIPAVGSPAPFTLTGDLGVISDQFRGKSVLTNIPPSVDTTPVC 60
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
DB 1 MAKTHFQGNBTATSGELPQVGDNLAEFNLVNTLGEVSSKDFQGRKLVNIPPSVDTTPVC 60
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
QY 61 ATSVRTFEDRAAA-SSATVLCVSKDLFPQAKRFGCAEGTENWMPASAFRDSFGEDYGV 119
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
DB 61 ATSVRKFNBAASLENTVLCISKDLFPALGRFCAGLEINVTVPASAFSTFGEDNGIVL 120
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
QY 120 ADGPAGLLARAIIVTGADGNVAYTELVPBIAQBPYEAALAL 163
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
DB 121 BGSPLAGLLARAIIVVDGQNKVLAHSELVIADEPNYAAALAGL 164
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||

RESULT 4
US-09-602-777A-172
; Sequence 172, Application US/09602777A
; Patent No. 6831165
; GENERAL INFORMATION:
; APPLICANT: Pompeius, Markus
; APPLICANT: Kroeger, Burkhard
; APPLICANT: Schroeder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Heberhauser, Gregor
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS
; TITLE OF INVENTION: INVOLVED IN HOMEOSTASIS AND ADAPATION
; FILE REFERENCE: BGI-128CP
; CURRENT APPLICATION NUMBER: US/09/602,777A
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 60/141031
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: DE 19931636.8
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19932125.6
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932126.4
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932127.2
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932128.0
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932129.9
; PRIOR FILING DATE: 1999-07-19
; PRIOR APPLICATION NUMBER: DE 19932226.0
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932920.6
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19932922.2
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19932924.9
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19932928.1
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19932930.3
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19932933.8
```

```

; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19932935.4
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19932973.7
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19933002.6
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19933003.4
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19933005.0
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19933006.9
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19941378.9
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: DE 19941379.7
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: DE 19941390.8
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: DE 19941391.6
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: DE 19942088.2
; PRIOR FILING DATE: 1999-09-03
; NUMBER OF SEQ ID NOS: 442
; SEQ ID NO 172
; LENGTH: 165
; TYPE: PR1
; ORGANISM: Corynebacterium glutamicum
US-09-602-777A-172

```

```

Query Match      56.1%; Score 466.5; DB 2; Length 165;
Best Local Similarity 56.7%; Pred. No. 1.4e-44;
Matches 93; Conservative 23; Mismatches 47; Indels 1; Gaps 1;

```

```

QY 1 MAQTTLGNAINTVGEIPAVGSPAPAFITLGGDLGVISDQFRGKSVLNIFFSVDTPTVC 60
    |||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 1 MATHTPOGNTATSGELPOVGDNLAEFNLVTELGKVSXDFQGRKLVNIFFSVDTPTVC 60
    |||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 61 ATSVRTFDEBAAA-SGATVLCVSKDLPPAOKRFGAGSTENWMPASAFRDSFGEDYGVTI 119
    |||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 61 ATSVRTFDEBAAA-SGATVLCVSKDLPPAOKRFGAGSTENWMPASAFRDSFGEDYGVTI 120
    |||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 120 ADGPMAGLLARAVIVIGADGVAAVTELVPEIAQEPNYEALAAAL 163
    |||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 121 EGSPPLKGLARAVIVIGADGVAAVTELVPEIAQEPNYEALAAAL 164
    |||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

```

```

RESULT 5
US-09-543-681A-4608
; Sequence 4608, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543.681A
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 4608
; LENGTH: 172
; TYPE: PR1
; ORGANISM: Proteus mirabilis
US-09-543-681A-4608

```

```

Query Match      53.1%; Score 441; DB 2; Length 172;
Best Local Similarity 55.2%; Pred. No. 1.1e-41;
Matches 90; Conservative 21; Mismatches 50; Indels 2; Gaps 2;
QY 3 QITLGNAINTVGELPAVGSPPAPAFITLGGDLGVISDQFRGKSVLNIFFSVDTPTVC 62
    |||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

```

```

DB 9 QVTIQQNAVTLIAGNPFVVGQKAADFSILVGDLADVSLAQFAGKRVKLTNIFFSVDTPTVC 68
QY 63 SVRTFDEBAAA-SGATVLCVSKDLPPAOKRFGAGSTENWMPASAFRDS-FGSDYGVTI 120
    |||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 69 SVRTFDEBAAA-SGATVLCVSKDLPPAOKRFGAGSTENWMPASAFRDS-FGSDYGVTI 128
    |||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 121 DGPMAAGLLARAVIVIGADGVAAVTELVPEIAQEPNYEALAAAL 163
    |||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 129 SGPLAGLSRAVIVIGADGVAAVTELVPEIAQEPNYEALAAAL 171
    |||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

```

```

RESULT 6
US-09-489-039A-7716
; Sequence 7716, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489.039A
; PRIOR FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 7716
; LENGTH: 176
; TYPE: PR1
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-7716

```

```

Query Match      50.4%; Score 418.5; DB 2; Length 176;
Best Local Similarity 53.0%; Pred. No. 3.9e-39;
Matches 89; Conservative 24; Mismatches 52; Indels 3; Gaps 3;

```

```

QY 1 MAQ-TITLGNAINTVGELPAVGSPPAPAFITLGGDLGVISDQFRGKSVLNIFFSVDTPTVC 59
    |||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 9 MSQTVHFQGNPVSVQGTIPQAGAKAQFTLVADLSVALSQYAGKRVKLTNIFFSVDTPTVC 68
    |||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 60 CATSVRTFDEBAAA-SGATVLCVSKDLPPAOKRFGAGSTENWMPASAFRDS-FGSDYGV 117
    |||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 69 CAASVRTFDEBAAA-SGATVLCVSKDLPPAOKRFGAGSTENWMPASAFRDS-FGSDYGV 128
    |||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 118 TIADGPMAGLLARAVIVIGADGVAAVTELVPEIAQEPNYEALAAALGA 165
    |||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 129 AIATGPLAGLAAVAVIVIGADGVAAVTELVPEIAQEPNYEALAAALGA 176
    |||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

```

```

RESULT 7
US-09-602-777A-174
; Sequence 174, Application US/09602777A
; Patent No. 6831165
; GENERAL INFORMATION:
; APPLICANT: Pompejus, Markus
; APPLICANT: Krcger, Burhard
; APPLICANT: Schroder, Hartwig
; APPLICANT: Zelder, Oskar
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS
; FILE REFERENCE: BGI-128CP
; CURRENT APPLICATION NUMBER: US/09/602.777A
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 60/141031
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: DE 19931636.8
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19932125.6
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932126.4
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932127.2
; PRIOR FILING DATE: 1999-07-09

```

```

PRIORITY APPLICATION NUMBER: US 60/085,131
PRIORITY FILING DATE: 1998-05-12
PRIORITY APPLICATION NUMBER: US 60/051,553
PRIORITY FILING DATE: 1997-07-02
NUMBER OF SEQ ID NOS: 5322
SEQ ID NO 3217
LENGTH: 172
TYPE: PRT
ORGANISM: Streptococcus pneumoniae
US-09-583-110-3217

Query Match          38.4%; Score 319.5; DB 2; Length 172;
Best Local Similarity 44.5%; Pred. No. 5,4e-28;
Matches 73; Conservative 23; Mismatches 63; Indels 5; Gaps 3;

Cy 4 ITLGNAINVTGELIPAVGSPAPAFITLTGDLGVISDQFRGKSVLNIIPSVDTVCATS 63
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 11 VTFLGNVPSFTGKQLQGDKALDFSLTTTDLSSKSLADFGKKVLSVPSIDTIGICSTQ 70
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Cy 64 VRTDER-AAASAGTVCYSKDLRFPAQRCCGAGETN-VMPASAFPSDGEVDGVITAD 121
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 71 TRRNERELAGIDNTVTVLTVSMDLRFPAQRCCGAGELDNLATLSDFPHSFGRDYALLNE 130
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Cy 122 GPMAGLIARAIIVYIGADGNVAAYTELVEPIAOEPPYEAALALGA 165
   ||||| : : : : : : : : : : : : : : : : : : : : : : : :
Db 131 W--HLIARAVFVLDNTNTIHVEYVDNINSEPFELAIIDAAKA 171
   ||||| : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 9
US-09-107-433-5126
Sequence 5126, Application US/09107433
Patent No. 6800744
GENERAL INFORMATION:
APPLICANT: Lynn A Doucette-Stamm and David Bush
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID
SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE
THERAPEUTICS
NUMBER OF SEQUENCES: 5206
CORRESPONDENCE ADDRESS:
ADDRESSEE: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: <Unknown>
OPERATING SYSTEM: <Unknown>
SOFTWARE: <Unknown>
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,433
FILING DATE: 30-Jun-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/085131
FILING DATE: May 12, 1998
APPLICATION NUMBER: 60/051553
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Arinello, Pamela Daneke
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-011
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 5126:
SEQUENCE CHARACTERISTICS:
LENGTH: 201 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYDROTHERMAL: YES
ORIGINAL SOURCE:

```



```

1 ORGANISM: Streptococcus pneumoniae
2 FEATURE:
3 NAME/KEY: misc_feature
4 LOCATION: (B) LOCATION 1..201
5 SEQUENCE DESCRIPTION: SEQ ID NO: 5126:
6 US-09-107-433-5126
7
8 Query Match      38.4%; Score 319.5; DB 2; Length 201;
9 Match Local Similarity 44.5%; Pred. No. 6.7e-28;
10 Matches 73; Conservative 23; Mismatches 63; Indels 5; Gaps 3;
11
12 QY      4 ITTGNKAINVTVGSELPAVGSPAPAFATTLTGADLGVISDPQFGSKSVLTINIPSVVTTPVCATS 63
13       :|::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||:
14 Db      40 VTRFENGVSTPGKQLQVGDKALDPSLTITLDLSKSLADPFGKKKVLVSVPISIDTIGCSNQ 99
15       ||::||::||::||::||::||::||::||::||::||::||::||::||::||::||:
16 QY      64 VRTFDER-AAASGATVLCVSKDLPEFAQRFCGARGTEN-VMPASAFRDSFGEDYGYTIAD 121
17       ||::||::||::||::||::||::||::||::||::||::||::||::||::||::||:
18 Db      100 TRFRNEBELAGDNTVVLTGSMWLEPFAQKRCGABGLDNALIMLSDFPHSGPRDYALLINE 159
19       ||::||::||::||::||::||::||::||::||::||::||::||::||::||::||:
20 QY      122 GPAGGLARAVLVIGADGNVAITYELVEIAOEPRIYEAAALGA 165
21       ||::||::||::||::||::||::||::||::||::||::||::||::||::||::||:
22 Db      160 W--HLLRAVFVLDITNTHIVEEYVDNIINSEPNFEALAAARA 200
23       ||::||::||::||::||::||::||::||::||::||::||::||::||::||::||:
24
25 RESULT 10
26 US-08-956-171B-5218
27 Sequence 5218, Application US/08956171B
28 Patent No. 6593114
29 GENERAL INFORMATION:
30 APPLICANT: Charles Kunach
31           Gil H. Choi
32           Patrick S. Dillon
33           Craig A. Rosen
34           Steven C. Barash
35           Michael R. Pannou
36 TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
37 NUMBER OF SEQUENCES: 5256
38 CORRESPONDENCE ADDRESS:
39 ADDRESSEE: Human Genome Sciences, Inc.
40 STREET: 9410 Key West Avenue
41 CITY: Rockville
42 STATE: Maryland
43 COUNTRY: USA
44 ZIP: 20850
45 COMPUTER READABLE FORM:
46 MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
47 COMPUTER: HP Vectra 486/33
48 OPERATING SYSTEM: MSDOS version 6.2
49 SOFTWARE: ASCII Text
50 CURRENT APPLICATION DATA:
51 APPLICATION NUMBER: US/08/956,171B
52 FILING DATE: 20-Oct-1997
53 CLASSIFICATION: <Unknown>
54 PRIOR APPLICATION DATA:
55 APPLICATION NUMBER: 60/009,861
56 FILING DATE: January 5, 1996
57 APPLICATION NUMBER: 08/781,986
58 FILING DATE: January 3, 1997
59 ATTORNEY/AGENT INFORMATION:
60 NAME: Mark J. Hyman
61 REGISTRATION NUMBER: 46,789
62 REFERENCE/DOCKET NUMBER: PB248P1
63 TELECOMMUNICATION INFORMATION:
64 TELEPHONE: (240) 314-1224
65 TELEFAX: (301) 309-8439
66 INFORMATION FOR SEQ ID NO: 5218:
67 SEQUENCE CHARACTERISTICS:
68 LENGTH: 173 amino acids
69 TYPE: amino acid
70 STRANDEDNESS: single
71 TOPOLOGY: linear
72 MOLECULE TYPE: protein
73 SEQUENCE DESCRIPTION: SEQ ID NO: 5218:

```

```

US-08-956-171E-5218

Query Match      37.8%, Score 314; DB 2; Length 173;
Best Local Similarity 42.3%, Pred. No. 2.3e-27;
Matches 69; Conservative 28; Mismatches 62; Indels 4; Gaps 2.

QY MAOITLRNAINTVGELPAVGSAPAFAFLTNGDLGVISDDPFRKGSVLNTLFPSTVPFVC 60
DB MTETTFPKGPFHLKGGQQTINBEDPAPDPFTVLNDNLNQVTLYADYAKKKLLISVPSIDTGC 69
QY 61 ATSVRTDERBAASAGATVLCVSXDLPPAOKRFGCAEGTENVMPASAFPD-SFGEDYGVTI 119
DB DOGRKRKFNSDASKESGIYLITISADLPFAQRKMCSAGLDNVITLSDRHDLSEFYENYGVVM 129
QY 120 ADGPMAGLARALVIVIGADSNVAATELVPEIAQBNRYAALAA 162
DB 130 EE---LRLRARAVFVLDAADNRVYKEIYSEGTDPEFDPAALAA 169

RESULT 11
US-08-781-986A-5218
Sequence 5218, Application US/08781986A
Patent No. 6737248
GENERAL INFORMATION:
Applicant: Charles Kunach
Title Of Invention: Staphylococcus aureus Polynucleotides and Sequences
Number Of Sequences: 5255
Correspondence Address:
Addressee: Human Genome Sciences, Inc.
Street: 9410 Key West Avenue
City: Rockville
State: Maryland
Country: USA
Zip: 20850
COMPUTER READABLE FORM:
Medium Type: Diskette, 3.50 inch, 1.4Mb storage
Computer: HP Vectra 486/33
Operating System: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
Application Number: US/08/781,986A
Filing Date:
Classification: 435
Prior Application Data:
Application Number:
Filing Date:
ATTORNEY/AGENT INFORMATION:
Name: Benson, Bob
Registration Number: 30,446
Reference/Docket Number: PP248PP
Telecommunication Information:
Telephone: (301) 309-8504
Telex: (301) 309-8512
Information For Seq ID NO: 5218:
SEQUENCE CHARACTERISTICS:
Length: 173 amino acids
Type: amino acid
Strandedness: single
Topology: linear
Molecule type: protein
US-08-781-986A-5218

Query Match      37.8%, Score 314; DB 2; Length 173;
Best Local Similarity 42.3%, Pred. No. 2.3e-27;
Matches 69; Conservative 28; Mismatches 62; Indels 4; Gaps 2.

QY 1 MAOITLRNAINTVGELPAVGSAPAFAFLTNGDLGVISDDPFRKGSVLNTLFPSTVPFVC 60
DB 10 MTETTFPKGPFHLKGGQQTINBEDPAPDPFTVLNDNLNQVTLYADYAKKKLLISVPSIDTGC 69
QY 61 ATSVRTDERBAASAGATVLCVSXDLPPAOKRFGCAEGTENVMPASAFPD-SFGEDYGVTI 119
DB 70 DOGRKRKFNSDASKESGIYLITISADLPFAQRKMCSAGLDNVITLSDRHDLSEFYENYGVVM 129
QY 120 ADGPMAGLARALVIVIGADSNVAATELVPEIAQBNRYAALAA 162
DB 130 EE---LRLRARAVFVLDAADNRVYKEIYSEGTDPEFDPAALAA 169

```

```
Oy      120 ADGPMAGLLARAIIVIGADGNVAYTELVPEIAQEPNTEAALAA 162
          |||||:|:|||:::|||
Db      130 EE--LRRLARAVFVLDAADNKVVKELVSEGTDFPDPDAAALAA 169
```

RESULT 12

```

US-09-710-279-252
/ Sequence 252, Application US/09710279
/ Patent No. 6703492
/ GENERAL INFORMATION:
/ APPLICANT: KIMBERLY, WILLIAM JOHN
/ TITLE OF INVENTION: STREPTOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
/ FILE REFERENCE: P03480US
/ CURRENT APPLICATION NUMBER: US/09/710,279
/ CURRENT FILING DATE: 2000-11-09
/ PRIOR APPLICATION NUMBER: 60/164,258
/ PRIOR FILING DATE: 1999-11-09
/ NUMBER OF SEQ ID NOS: 4472
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 252
/ LENGTH: 164
/ TYPE: PRN
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence: synthetic
/ US-09-710-279-252

```

Query Match	36.9%	Score 307	DB 26	Length 164
Best Local Similarity	40.5%	Pred. No. 1.3e-26		
Matches 66	Conservative 29	Mismatches 64	Indels 4	Gaps 2

QY 1 MAQITLRGAINVVGELPAVGGPAPAFETLTGGDLGVISDDPGRGSLVLTINFSPVDVPC 60

Db 1 MTQITTKNPNPKIISGSEVNEBGDIAPAPFVTLVDNLSINQITLDDYKNNKKLISVPIBIDGVC 60

QY 61 ATSVRTFDEBAAASGATVLCVSKDLPFAQKRCFGAEGTEINWPPASAFPD-SFGEDYGVTI 119

Db 61 DSGTRRFKNEBASABDGVLLTISVDLPFAQKRCQASGLDNYTTTSDHKOLSRGNNTGLVM 120

QY 120 ADGPMAGLLAPAAIVIVIGADGNVAYTELVBEIQBPYEAALAA 162

Db 121 DE---LRLRLARSVFVLTENNNKVVKEIYSEGVNTYDPDFEALKA 160

RESULT 13

```

US-09-134-001C-3764
; Sequence 3764, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134, 001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 3764
; LENGTH: 168
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
; US-09-134-001C-3764

```

Query Match	36.9%	Score 307	DB 2	Length 168
Best Local Similarity	40.5%	Pred. No. 1.3e-26		
Matches 66	Conservative 29	Mismatches 64	Indels 4	Gaps 2
QY	1	MAQTTIRNAINVYVGELEPAVSGPAPAFUTLTGGDLGVSSDDQKRGKSVLNTPPSYDTVC	60	

Db

5 MTQITFKNNBIKLSGSEVNEGDIAPNFTVLNDSLNQITLDDYKNKKCLISVIPISIDTGC 64

Dy 61 ATSVRTFDERAALSGATVLCVKSLPFAQKRFCCGAGTENMPASAFRD-SGEDYCVTI 119
 : | : | : | : | : | : | : | : | : | : | : | : | :
Db 65 DSGTRKFNEASAEADGVLTITISVDLPFAQRWCASGLDNVITLSPHKDLSEGRNYGLVM 124

```

QY      120 ADGPMAGLLARAIIVIGADGNVATTELPELAEQENYEALAA 162
      : |||::: | : | : | : ||| |
DB      125 DE---LRLLARSVFVLNENNNKVYKEIVSEGTNYPDFEALKA 164

```

RESULT 14

```

US-09-902-540-15283
; Sequence 15283, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 15283
; LENGTH: 171
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
; US-09-902-540-15283

```

Query Match	36.3%	Score 302	DB 2	Length 171
Best Local Similarity	40.6%	Pred. No. 5e-26		
Matches 67	Conservative 30	Mismatches 62	Indels 6	Gaps 3

QY 4 ITLRGAAIYVWSELPAVCSAPAFALITGGDGLGISDDFRKSVLILNFPVDIVPVCATS 63
::
Db 8 VTFKGQPTVLTVGDEYKVGDPAEDFTVFGLNDAYRLSDPKVSVVVVSAAPSVDTRVCQAQ 67

QY 64 VRTFDSRAAASG--ATVLCTSKDLPPAQKRCGAHGTENNMPAAPRD-SGEIDYGYTIA 120Q
::
Db 68 LRFNFNEAALADPDVQWWEVTLDLPFALGRFLGAGIONVTTLSDYDRREGEKGIQYMK 127Q

```

Qy      121 DGPAGLLARAIIVIGADGNVAYTELVPBIAQEPNRYEALLALGA 165
      : ||||| : : : : ||| : ||| : : : |
Db      128 E---LGLLARSTFVVDREGKVFREIYVBMTHBPDYDGAMKAVRA 169

```

RESULT 15

```

1  US-08-961-083-44
2  ; Sequence 44, Application US/08961083
3  ; Patent No. 6159469
4  ;
5  ; GENERAL INFORMATION:
6  ;
7  ; APPLICANT: Choi et. al.
8  ;
9  ; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
10 ;
11 ; NUMBER OF SEQUENCES: 452
12 ;
13 ; CORRESPONDENCE ADDRESS:
14 ; ADDRESSEE: Human Genome Sciences, Inc.
15 ;
16 ; STREET: 9410 Key West Avenue
17 ; CITY: Rockville
18 ; STATE: Maryland
19 ;
20 ; COUNTRY: USA
21 ;
22 ; ZIP: 20850
23 ;
24 ; COMPUTER READABLE FORM:
25 ;
26 ; MEDIUM TYPE: Diskette, 3.50 inch, 1.44b storage
27 ;
28 ; COMPUTER: HP Vectra 486/33
29 ; OPERATING SYSTEM: MSDOS version 6.2
30 ;
31 ; SOFTWARE: ASCII Text
32 ;
33 ; CURRENT APPLICATION DATA:
34 ;
35 ; APPLICATION NUMBER: US/08/961,083
36 ;
37 ;
38 ;
39 ;
40 ;
41 ;
42 ;
43 ;
44 ;
45 ;
46 ;
47 ;
48 ;
49 ;
50 ;
51 ;
52 ;
53 ;
54 ;
55 ;
56 ;
57 ;
58 ;
59 ;
60 ;
61 ;
62 ;
63 ;
64 ;
65 ;
66 ;
67 ;
68 ;
69 ;
70 ;
71 ;
72 ;
73 ;
74 ;
75 ;
76 ;
77 ;
78 ;
79 ;
80 ;
81 ;
82 ;
83 ;
84 ;
85 ;
86 ;
87 ;
88 ;
89 ;
90 ;
91 ;
92 ;
93 ;
94 ;
95 ;
96 ;
97 ;
98 ;
99 ;
100 ;

```

```

: FILING DATE:
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER:
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Brookes, A. Anders
: REGISTRATION NUMBER: 36,373
: REFERENCE/DOCKET NUMBER: PB340P2
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (301) 309-8504
: TELEFAX: (301) 309-8512
: INFORMATION FOR SEQ ID NO: 44:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 144 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-08-961-083-44

```

```

Query Match          34.0%; Score 282.5; DB 2; Length 144;
Best Local Similarity 44.8%; Pred. No. 6,2e-24;
Matches 65; Conservative 19; Mismatches 56; Indels 5; Gaps 3;

```

```

QY 15 GELPVGSPAPAFTLTGSDLGVISDDQFRGKSVLNIFPSVDTPVCATSVTFPDER-AAA 73
   | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 3 GKQLQVGDKALDPSLTITTDLSKKSADFDGKKVLSVPSIDTIGCTQTRRFNEELAGL 62
   | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 74 SGATVLCVSKDLPFAQRCFGAESTEN-VMPASAFRDSFGEDYGVTTADGPMAGLLARAI 132
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 63 DNTVVLTVSMDLPPAQRCRCGAEGLDVAIMLSDYFDHSFGSDYALLINEM---HLLARAV 119
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 133 VVIGADGNVAYTELYPEEIAQEPNVE 157
   | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 120 FVLDTDTIRRYEYVDNINSEPNFE 144
   | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

Search completed: April 14, 2006, 17:37:50
 Job time : 16.2571 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: April 14, 2006, 18:27:13 ; Search time 57.9026 Seconds
(without alignments)
1190.652 Million cell updates/sec

Title: US-10-620-246-8

Perfect score: 831
Sequence: 1 MAQITLRGNAINTVGSLPAV.....LVEBIAOBPNYEAALALGA 165

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA.Main:
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	831	100.0	165	3	US-09-791-171-8
2	831	100.0	165	4	US-10-620-246-8
3	831	100.0	165	4	US-10-332-512A-9
4	821	98.8	165	3	US-09-804-980-8
5	473.5	57.0	165	3	US-09-738-626-4707
6	417.5	50.2	168	5	US-10-893-671-45
7	320.5	38.6	167	5	US-10-472-928-3396
8	319.5	38.4	201	5	US-10-617-320-5126
9	314	37.8	173	2	US-08-781-986A-5218
10	314	37.8	173	4	US-10-329-624-5218
11	307	36.9	168	4	US-10-724-972A-5919
12	282.5	34.0	144	3	US-09-765-372-44
13	282.5	34.0	144	6	US-11-106-649-44
14	254	30.6	166	4	US-10-335-977-7878
15	222.5	26.8	139	4	US-10-335-977-7877
16	195.5	23.5	108	3	US-09-939-980-443
17	138.5	16.7	164	4	US-10-369-493-22947
18	138.5	16.7	164	5	US-10-732-923-21693
19	136.5	16.4	155	5	US-10-732-923-20983
20	135.5	16.3	148	5	US-10-732-923-21601
21	135.5	16.3	156	5	US-10-732-923-21643
22	135.5	16.3	157	5	US-10-732-923-21645
23	133.5	16.0	157	5	US-10-732-923-20994
24	133.5	16.0	151	5	US-10-732-923-21646
25	131.5	15.8	157	5	US-10-732-923-21666
26	131.5	15.8	159	5	US-10-732-923-21673
27	130.5	15.7	224	5	US-10-732-923-21442

28	130	15.6	155	5	US-10-732-923-21569	Sequence 21569, A
29	129.5	15.6	257	5	US-10-732-923-21110	Sequence 21310, A
30	129	15.5	157	5	US-10-732-923-21660	Sequence 21660, A
31	128.5	15.5	155	5	US-10-732-923-21588	Sequence 21588, A
32	125.5	15.1	132	5	US-10-732-923-21240	Sequence 21240, A
33	125.5	15.1	195	5	US-10-732-923-21244	Sequence 21244, A
34	125	15.0	185	5	US-10-732-923-21421	Sequence 21421, A
35	124.5	15.0	185	5	US-10-732-923-21238	Sequence 21238, A
36	124	14.9	155	5	US-10-732-923-21599	Sequence 21599, A
37	123.5	14.9	171	5	US-10-732-923-21696	Sequence 21696, A
38	122.5	14.7	145	5	US-10-732-923-21606	Sequence 21606, A
39	122.5	14.7	159	5	US-10-732-923-21633	Sequence 21633, A
40	122.5	14.7	163	5	US-10-732-923-21633	Sequence 21633, A
41	122.5	14.7	195	5	US-10-732-923-21634	Sequence 21634, A
42	122.5	14.7	228	5	US-10-732-923-20973	Sequence 20973, A
43	122.5	14.7	576	5	US-10-732-923-21242	Sequence 21242, A
44	122.5	14.6	198	5	US-10-732-923-20953	Sequence 20953, A
45	121.5	14.6				

ALIGNMENTS

```
RESULT 1
US-09-791-171-8
; Sequence 8, Application US/09791171
; Patent No. US2002009436A1
; GENERAL INFORMATION:
; APPLICANT: ANDERSEN, Peter
; APPLICANT: NIELSEN, Rikke
; APPLICANT: OSTTINGER, Thomas
; APPLICANT: RASMUSSEN, Peter Birk
; APPLICANT: ROSENKRANDS, Ida
; APPLICANT: WELDKINGH, Karin
; APPLICANT: FIORIO, Walter
; TITLE OF INVENTION: NUCLEIC ACIDS FRAGMENTS AND POLYPEPTIDE FRAGMENTS
; FILE REFERENCE: 670001-2002.1
; CURRENT APPLICATION NUMBER: US/09/791,171
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 09/050,739
; PRIOR FILING DATE: 1998-03-30
; PRIOR APPLICATION NUMBER: 0376/97
; PRIOR FILING DATE: 1997-04-02
; PRIOR APPLICATION NUMBER: 1277/97
; PRIOR FILING DATE: 1997-11-10
; PRIOR APPLICATION NUMBER: 60/044,624
; PRIOR FILING DATE: 1997-04-18
; PRIOR APPLICATION NUMBER: 60/070,488
; PRIOR FILING DATE: 1998-01-05
; NUMBER OF SEQ ID NOS: 173
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 165
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-09-791-171-8
Query Match 100.0%; Score 831; DB 3; Length 165;
Best Local Similarity 100.0%; Pred. No. 2.4e-79;
Matches 165; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY	1	MAQITLRGNAINTVGSLPAVPTLTGGDLGVISDDPRGKSVLNTIPSVDTTPVC	60
DB	1	MAQITLRGNAINTVGSLPAVGPAPAFITLGGDLGVISDDPRGKSVLNTIPSVDTTPVC	60
QY	61	ATSVRTFDRRAAASGATVLCVSKDLPFAQKRFCAAGTEWVMPASAFRDSFGEDYGVTTA	120
DB	61	ATSVRTFDRRAAASGATVLCVSKDLPFAQKRFCAAGTEWVMPASAFRDSFGEDYGVTTA	120
QY	121	DGPMAGLARAIYVIGADGNVATBLVPRIAOBPNYEAALALGA	165
DB	121	DGPMAGLARAIYVIGADGNVATBLVPRIAOBPNYEAALALGA	165

```
RESULT 2
US-10-620-246-8
; Sequence 8, Application US/10620246
; Publication No. US20040115211A1
; GENERAL INFORMATION:
; APPLICANT: ANDERSEN, Peter
; APPLICANT: NIELSEN, Rikke
; APPLICANT: OETTINGER, Thomas
; APPLICANT: RASMUSSEN, Peter Birk
; APPLICANT: ROSENKRANDS, Ida
; APPLICANT: WELDINGH, Karin
; APPLICANT: FLORIO, Walter
; TITLE OF INVENTION: NUCLEIC ACIDS FRAGMENTS AND POLYPEPTIDE FRAGMENTS
; TITLE OF INVENTION: DERIVED FROM M. TUBERCULOSIS
; FILE REFERENCE: 670001-2002.1A
; CURRENT APPLICATION NUMBER: US/10/620,246
; PRIOR FILING DATE: 2003-07-15
; PRIOR APPLICATION NUMBER: 09/050,739
; PRIOR FILING DATE: 1998-03-30
; PRIOR APPLICATION NUMBER: 0376/97
; PRIOR FILING DATE: 1997-04-02
; PRIOR APPLICATION NUMBER: 1277/97
; PRIOR FILING DATE: 1997-11-10
; PRIOR APPLICATION NUMBER: 60/044,624
; PRIOR FILING DATE: 1997-04-18
; PRIOR APPLICATION NUMBER: 60/070,488
; PRIOR FILING DATE: 1998-01-05
; PRIOR APPLICATION NUMBER: 10/138,473
; PRIOR FILING DATE: 2002-05-02
; PRIOR APPLICATION NUMBER: 09/791,171
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 09/415,884
; PRIOR FILING DATE: 1999-10-08
; PRIOR APPLICATION NUMBER: 60/116,673
; PRIOR FILING DATE: 1999-01-21
; PRIOR APPLICATION NUMBER: 1281/98
; PRIOR FILING DATE: 1998-10-08
; NUMBER OF SEQ ID NOS: 173
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 165
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-10-620-246-8

Query Match      100.0%; Score 831; DB 4; Length 165;
Best Local Similarity 100.0%; Pred. No. 2.4e-79;
Matches 165; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAQITLRGNAINTVGEHPVAGSPAPAFITLTGGDLGVYISSDQFRGKSVLNTIPPSVDTFPC 60
DB 1 MAQITLRGNAINTVGEHPVAGSPAPAFITLTGGDLGVYISSDQFRGKSVLNTIPPSVDTFPC 60
QY 61 ATSVRTFDEARRAASGATVLCVSKDLPFAQKRCGAGTENWMPASAFRDSFGEDYGVTTA 120
DB 61 ATSVRTFDEARRAASGATVLCVSKDLPFAQKRCGAGTENWMPASAFRDSFGEDYGVTTA 120
QY 121 DGPMAGLARAIIVIGADGNVATYELVPEIAQEPNYEALALGA 165
DB 121 DGPMAGLARAIIVIGADGNVATYELVPEIAQEPNYEALALGA 165

RESULT 3
US-10-332-512A-9
; Sequence 9, Application US/10332512A
; Publication No. US20040180056A1
; GENERAL INFORMATION:
; APPLICANT: ORME, Ian M.
; APPLICANT: BELISLE, John T.
; TITLE OF INVENTION: MID-LIFE VACCINE AND METHODS FOR BOOSTING ANTI-MYCOBACTERIAL IMM
; FILE REFERENCE: 38861-186292
```

```
; CURRENT APPLICATION NUMBER: US/10/332,512A
; CURRENT FILING DATE: 2003-01-10
; PRIOR APPLICATION NUMBER: PCT/US01/21717
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/217,646
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
; LENGTH: 165
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-10-332-512A-9

Query Match      100.0%; Score 831; DB 4; Length 165;
Best Local Similarity 100.0%; Pred. No. 2.4e-79;
Matches 165; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAQITLRGNAINTVGEHPVAGSPAPAFITLTGGDLGVYISSDQFRGKSVLNTIPPSVDTFPC 60
DB 1 MAQITLRGNAINTVGEHPVAGSPAPAFITLTGGDLGVYISSDQFRGKSVLNTIPPSVDTFPC 60
QY 61 ATSVRTFDEARRAASGATVLCVSKDLPFAQKRCGAGTENWMPASAFRDSFGEDYGVTTA 120
DB 61 ATSVRTFDEARRAASGATVLCVSKDLPFAQKRCGAGTENWMPASAFRDSFGEDYGVTTA 120
QY 121 DGPMAGLARAIIVIGADGNVATYELVPEIAQEPNYEALALGA 165
DB 121 DGPMAGLARAIIVIGADGNVATYELVPEIAQEPNYEALALGA 165

RESULT 4
US-09-804-980-8
; Sequence 8, Application US/09804980
; Publication No. US20030147697A1
; GENERAL INFORMATION:
; APPLICANT: Statens Serum Institut
; APPLICANT: Anderson, Peter
; TITLE OF INVENTION: M. Tuberculosis Antigens
; FILE REFERENCE: 670001-2002.4
; CURRENT APPLICATION NUMBER: US/09/804,980
; CURRENT FILING DATE: 2001-03-12
; NUMBER OF SEQ ID NOS: 257
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 8
; LENGTH: 165
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-09-804-980-8

Query Match      98.8%; Score 821; DB 3; Length 165;
Best Local Similarity 98.8%; Pred. No. 2.7e-78;
Matches 163; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAQITLRGNAINTVGEHPVAGSPAPAFITLTGGDLGVYISSDQFRGKSVLNTIPPSVDTFPC 60
DB 1 MAQITLRGNAINTVGEHPVAGSPAPAFITLTGGDLGVYISSDQFRGKSVLNTIPPSVDTFPC 60
QY 61 ATSVRTFDEARRAASGATVLCVSKDLPFAQKRCGAGTENWMPASAFRDSFGEDYGVTTA 120
DB 61 ATSVRTFDEARRAASGATVLCVSKDLPFAQKRCGAGTENWMPASAFRDSFGEDYGVTTA 120
QY 121 DGPMAGLARAIIVIGADGNVATYELVPEIAQEPNYEALALGA 165
DB 121 DGPMAGLARAIIVIGADGNVATYELVPEIAQEPNYEALALGA 165

RESULT 5
US-09-738-626-4707
; Sequence 4707, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
```

APPLICANT: MIZOGUCHI, HIROSHI
APPLICANT: ANDO, SEIRO
APPLICANT: HAYASHI, MIKIRO
APPLICANT: OCHIAI, KEIKO
APPLICANT: YOKOI, HARUHIKO
APPLICANT: TATEISHI, NAOKO
APPLICANT: SENOH, AKIHIRO
APPLICANT: IKEDA, MASATO
APPLICANT: OZAKI, AKIO
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-125
CURRENT APPLICATION NUMBER: US/09/738,626
CURRENT FILING DATE: 2000-12-18
PRIOR APPLICATION NUMBER: JP 99/377484
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: JP 00/159162
PRIOR FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: JP 00/280988
PRIOR FILING DATE: 2000-08-03
NUMBER OF SEQ ID NOS: 7059
SOFTWARE: PatentIn ver. 3.0
SEQ ID NO 4707
LENGTH: 165
TYPE: PRT
ORGANISM: Corynebacterium glutamicum
US-09-738-626-4707

Query Match 57.0%; Score 473.5; DB 3; Length 165;
Best Local Similarity 57.3%; Pred. No. 1.2e-41;
Matches 94; Conservative 24; Mismatches 45; Indels 1; Gaps 1;

QY 1 MAQITLGNAINTVGELPAVGSPPAPFTLTGDLGVISSDQFRGKSVLTNIFFSVDPVC 60
DB 1 MAKHPQGENRATSGELPQVGDNLAEFVLVNTBELGVSSKDFQGRKVLNIFFSVDTVC 60
QY 61 ATSVRTDERAAA-SGATVLCVSKDLPFAQRFCGAGSTENVMPASAFRDSFGEDYGTI 119
DB 61 ATSVRKRNKAASLNTVLCISKDLPPALGRFCGABIEIENVTPVSARSTFGEDMGIVL 120
QY 120 ADGMAGLLARAIVVIGADGNVATLVEPIAQBPNYEAALAL 163
DB 121 EGSPKGLMARSIVVDENGKVAATQVLDVDEISTEDYDALAL 164

RESULT 6

US-10-693-671-45
Sequence 45, Application US/10893671
Publication No. US20050064527A1
GENERAL INFORMATION:
APPLICANT: Levy, Stuart, et. al.
TITLE OF INVENTION: NMR COMPOSITIONS AND THEIR METHODS OF USE
FILE REFERENCE: PRZ-043
CURRENT APPLICATION NUMBER: US/10/893,671
CURRENT FILING DATE: 2004-07-15
PRIOR APPLICATION NUMBER: US/09/801,563
PRIOR FILING DATE: 2001-03-08
PRIOR APPLICATION NUMBER: 60/188,362
PRIOR FILING DATE: 2000-03-10
NUMBER OF SEQ ID NOS: 98
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 45
LENGTH: 168
TYPE: PRT
ORGANISM: Escherichia coli
FEATURES:
OTHER INFORMATION: complement of position 6467-6937 of seq id 44
US-10-693-671-45

Query Match 50.2%; Score 417.5; DB 5; Length 168;
Best Local Similarity 52.4%; Pred. No. 1e-35; Indels 3; Gaps 3;
Matches 88; Conservative 23; Mismatches 54; Indels 3; Gaps 3;
QY 1 MAQ-ITLGNAINTVGELPAVGSPPAPFTLTGDLGVISSDQFRGKSVLTNIFFSVDPVC 59

DB 1 MSQTVHFGNVTYANSTPQAGSKAQFTTLVAKLSDVTLQCFAGKRVLTNIFFSVDTVC 60
QY 60 CATSVRTDERAAA-ASGATVLCVSKDLPFAQRFCGAGSTENVMPASAFRDS-FGEDYGV 117
DB 61 CAAVRKRNKAASLNTVLCISADLPFAQRFCGABIEIENVTPVSARSTFGEDMGIVL 120
QY 118 TIADGPMAGLLARAIVVIGADGNVATLVEPIAQBPNYEAALALGA 165
DB 121 AIDQPKGLMARSIVVDENDVILFSLVDVETITPEPYEAALAVLKA 168

RESULT 7

US-10-472-928-3396
Sequence 3396, Application US/10472928
Publication No. US2005020813A1
GENERAL INFORMATION:
APPLICANT: CHIRON SPA
TITLE OF INVENTION: THE INSTITUTE FOR GENOMIC RESEARCH
FILE REFERENCE: P026926WO
CURRENT APPLICATION NUMBER: US/10/472,928
CURRENT FILING DATE: 2003-09-26
PRIOR APPLICATION NUMBER: GB-0107658.7
PRIOR FILING DATE: 2001-03-27
NUMBER OF SEQ ID NOS: 4979
SOFTWARE: SeqMan99, version 1.03
SEQ ID NO 3396
LENGTH: 167
TYPE: PRT
ORGANISM: Streptococcus pneumoniae
FEATURES:
OTHER INFORMATION: chiol peroxidase (psad)
OTHER INFORMATION: cellular location: cytoplasm
OTHER INFORMATION: Similar to strain R6 sequence 15903538 (1.B-96)
US-10-472-928-3396

Query Match 38.6%; Score 320.5; DB 5; Length 167;
Best Local Similarity 44.5%; Pred. No. 1.8e-25;
Matches 73; Conservative 23; Mismatches 63; Indels 5; Gaps 3;

QY 4 ITLGNAINTVGELPAVGSPPAPFTLTGDLGVISSDQFRGKSVLTNIFFSVDPVCAT 63
DB 6 ITLGNAINTVGELPAVGSPPAPFTLTGDLGVISSDQFRGKSVLTNIFFSVDPVCAT 63
QY 64 VRTFDER-AAASGATVLCVSKDLPFAQRFCGAGSTENVMPASAFRDSFGEDYGTAD 121
DB 66 TRRFNEBLAGDNTVVLTVSMDFPFAQRFCGABIEIENVTPVSARSTFGEDMGIVL 125
QY 122 GPMAGLLARAIVVIGADGNVATLVEPIAQBPNYEAALALGA 165
DB 126 W--HLLARAVFVLDTNTIRYVYVNDINSEPIFEALAAKA 166

RESULT 8

US-10-617-320-5126
Sequence 5126, Application US/10617320
Publication No. US20050136404A1
GENERAL INFORMATION:
APPLICANT: Lynn A Doucette-Stamm and David Bush
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID
SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE
THERAPEUTICS
NUMBER OF SEQUENCES: 5206

CORRESPONDENCE ADDRESS:
ADDRESSER: GENOMI THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660

```

COMPUTER: <Unknown>
OPERATING SYSTEM: <Unknown>
SOFTWARE: <Unknown>
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/617,320
FILING DATE: 10-Jul-2003
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/107,433
FILING DATE: 30-Jun-1998
APPLICATION NUMBER: 60/085131
FILING DATE: May 12, 1998
APPLICATION NUMBER: 60/051553
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Arinello, Pamela Deneke
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-011
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 5126:
SEQUENCE CHARACTERISTICS:
LENGTH: 201 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: Streptococcus pneumoniae
FEATURE:
NAME/KEY: misc feature
LOCATION: (B) LOCATION 1...201
SEQUENCE DESCRIPTION: SEQ ID NO: 5126:
US-10-617-320-5126

```

```

Query Match      38.4%; Score 319.5; DB 5; Length 201;
Best Local Similarity 44.5%; Pred. No. 2.9e-25;
Matches 73; Conservative 23; Mismatches 63; Indels 5; Gaps 3;

```

```

QY 4 ITLRGNAINTVGELPAVGSPPAPFTLGGDLGVISDQFRGKSVLTINIPSVDTFVCATS 63
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
DB 40 VTFRGNVSTPTGKQLOVGRKALDPSLTITDLSKSLADFDGKKVLSVPSIDIGICSTQ 99
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
QY 64 VRTFDER-AAASGATVLCVSKDLPFAQKRFCAEGTENVMPASAFDSFGEDYGVTTAD 121
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
DB 100 TRFRNBLAGLDNTVVLTVSWDLPFAQKRWCAAGLDNATLSDYFPHSFRDYALLINE 159
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
QY 122 GPMAGLARAIIVYIGADGNVAYTELPPIAQPENYEAALAA 165
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
DB 160 W--HLIARAIVFLDNTIHYVYVDNINSEPNFEAAIAAKA 200
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||

```

```

RESULT 9
US-08-781-986A-5218
Sequence 5218, Application US/08781986A
Publication No. US20030054436A1
GENERAL INFORMATION:
APPLICANT: Charles Kunach
TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
NUMBER OF SEQUENCES: 5255
CORRESPONDENCE ADDRESS:
ADDRESSER: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:

```

```

APPLICATION NUMBER: US/08/781,986A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Benson, Bob
REGISTRATION NUMBER: 30,446
REFERENCE/DOCKET NUMBER: PB248PP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 5218:
SEQUENCE CHARACTERISTICS:
LENGTH: 173 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-781-986A-5218

```

```

Query Match      37.8%; Score 314; DB 2; Length 173;
Best Local Similarity 42.3%; Pred. No. 8.9e-25;
Matches 69; Conservative 28; Mismatches 62; Indels 4; Gaps 2;

```

```

QY 1 MAQITLRGNAINTVGELPAVGSPPAPFTLGGDLGVISDQFRGKSVLTINIPSVDTFVC 60
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
DB 10 MTEITFRKGPRIHLKGGQINSGDFAPDPTVLDNDLNGVTLADYAGKKGLISVPSIDTGV 69
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
QY 61 ATSVRTDERAAASGATVLCVSKDLPFAQKRFCAEGTENVMPASAFD-SFGEDYGVTT 119
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
DB 70 DOQTRKNSDASKEEGVLTISADLPFAQKRWCAAGLDNATLSDYFPHSFRDYALLINE 129
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
QY 120 ADGPMAGLARAIIVYIGADGNVAYTELPPIAQPENYEAALAA 162
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
DB 130 EE--LRLIARAIVFLDADNKKVYKIVSGETPFPPDALLAA 169
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||

```

```

RESULT 10
US-10-329-624-5218
Sequence 5218, Application US/10329624
Publication No. US20040043037A1
GENERAL INFORMATION:
APPLICANT: Charles Kunach
GIL H. Choi
Patrick S. Dillon
Craig A. Rosen
Steven C. Barash
Michael R. Fannon
TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
NUMBER OF SEQUENCES: 5256
CORRESPONDENCE ADDRESS:
ADDRESSER: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/329,624
FILING DATE: 27-Dec-2002
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/956,171
FILING DATE: October 20, 1997
APPLICATION NUMBER: 60/009,861
FILING DATE: January 5, 1996
APPLICATION NUMBER: 08/781,986

```


FILING DATE: January 3, 1997
 ATTORNEY/AGENT INFORMATION:
 NAME: Mark J. Hyman
 REGISTRATION NUMBER: 46,789
 REFERENCE/DOCKET NUMBER: PB248P1D1
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (240) 314-1224
 TELEFAX: (301) 309-8439
 INFORMATION FOR SEQ ID NO: 5218:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 173 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 SEQUENCE DESCRIPTION: SEQ ID NO: 5218:
 US-10-329-624-5218

Query Match 37.8%; Score 314; DB 4; Length 173;
 Best Local Similarity 42.3%; Pred. No. 8.9e-25;
 Matches 69; Conservative 28; Mismatches 62; Indels 4; Gaps 2;

QY 1 MAQITLGNAINTVGELPAVSGPAPAFTLTGSDLGVISSDQFRKSVLNIIFPSVDTPVC 60
 DB 10 MTEITFGGPIHLKGGQINSGDPAPDFTVLDNDLNQVTLADYAKKKLISVPSIDYVC 69
 QY 61 ATSVRTDERAASGATVLCVSKLPFAQKRFCAEGTENVPASAFRD-SFGSDYGVTI 119
 DB 70 DQGRKNSDASKEGIVLTISADLPFAQKRCASAGLDNVTISDRHDSFGNRYGVIM 129
 QY 120 ADGPMAGLARAIVIGADGNVATLVPETIAOEPNTEALAA 162
 DB 130 EE--LRLARAVFLVLDADNVVYKEIVSSEGTDPDPDAALAA 169

RESULT 11
 US-10-724-972A-5919
 Sequence 5919, Application US/10724972A
 Publication No. US2004014734A1
 GENERAL INFORMATION:
 APPLICANT: Doucette-Stamm, Lynn
 APPLICANT: Bush, David
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
 TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
 FILE REFERENCE: PAT03-16
 CURRENT APPLICATION NUMBER: US/10/724,972A
 CURRENT FILING DATE: 2003-12-01
 PRIOR APPLICATION NUMBER: 09/450,969
 PRIOR FILING DATE: 1999-11-29
 PRIOR APPLICATION NUMBER: 09/134,001
 PRIOR FILING DATE: 1998-08-13
 PRIOR APPLICATION NUMBER: 60/064,964
 PRIOR FILING DATE: 1997-11-08
 PRIOR APPLICATION NUMBER: 60/055,779
 PRIOR FILING DATE: 1997-08-14
 NUMBER OF SEQ ID NOS: 7544
 SEQ ID NO 5919
 LENGTH: 168
 TYPE: PRT
 ORGANISM: S.epidermidis
 US-10-724-972A-5919

Query Match 36.9%; Score 307; DB 4; Length 168;
 Best Local Similarity 40.5%; Pred. No. 4.7e-24;
 Matches 66; Conservative 29; Mismatches 64; Indels 4; Gaps 2;

QY 1 MAQITLGNAINTVGELPAVSGPAPAFTLTGSDLGVISSDQFRKSVLNIIFPSVDTPVC 60
 DB 5 MTQITFGNPIKISGSEVNEGDIAPNFTVLDNSLNQITLDDYKXKKGLISVPSIDYVC 64
 QY 61 ATSVRTDERAASGATVLCVSKLPFAQKRFCAEGTENVPASAFRD-SFGSDYGVTI 119
 DB 65 DSQTRKNEBASADGVLTISVDLPFAQKRCASAGLDNVTISDRHDSFGNRYGVIM 124

QY 120 ADGPMAGLARAIVIGADGNVATLVPETIAOEPNTEALAA 162
 DB 125 DE--LRLARAVFLVLDADNVVYKEIVSSEGTDPDPDAALAA 164

RESULT 12
 US-09-765-272-44
 Sequence 44, Application US/09765272
 Patent No. US20020061545A1
 GENERAL INFORMATION:
 APPLICANT: Choi et al.
 TITLE OF INVENTION: Streptococcus pneumoniae Antigen and Vaccines
 NUMBER OF SEQUENCES: 452
 CORRESPONDENCE ADDRESS:
 ADDRESSER: Human Genome Sciences, Inc.
 STREET: 9410 Key West Avenue
 CITY: Rockville
 STATE: Maryland
 COUNTRY: USA
 ZIP: 20850
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
 COMPUTER: HP Vectra 486/33
 OPERATING SYSTEM: MSDOS version 6.2
 SOFTWARE: ASCII Text
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/765,272
 FILING DATE: 22-Jan-2001
 CLASSIFICATION: <Unknown>
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/961,083
 FILING DATE: <Unknown>
 ATTORNEY/AGENT INFORMATION:
 NAME: Brooke, A. Anders
 REGISTRATION NUMBER: 36,373
 REFERENCE/DOCKET NUMBER: PB340P2
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (301) 309-8504
 TELEFAX: (301) 309-8512
 INFORMATION FOR SEQ ID NO: 44:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 144 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 SEQUENCE DESCRIPTION: SEQ ID NO: 44:
 US-09-765-272-44

Query Match 34.0%; Score 282.5; DB 3; Length 144;
 Best Local Similarity 44.8%; Pred. No. 1.5e-21;
 Matches 65; Conservative 19; Mismatches 56; Indels 5; Gaps 3;
 QY 15 GELPAVSGPAPAFTLTGSDLGVISSDQFRKSVLNIIFPSVDTPVCATSVRTDER-AAA 73
 DB 3 GKQLQVGDKALDPSLTITLDSKSLADFDGKKVLSVPSIDYGCSTQRRENEBLAGL 62
 QY 74 SGATVLCVSKOLPFAQKRFCAEGTEN-VMPASAFRDSFREDYGVTIADPMAGLARAI 132
 DB 63 DNTVLTIVSMDLPPAQRKWCAGAGLDNATLSDYFHSFGDYALLINW---HLARAV 119
 QY 133 VVIGADGNVATLVPETIAOEPNTE 157
 DB 120 FVLDTNTVIRYEVVDNINSEPNFE 144

RESULT 13
 US-11-106-649-44
 Sequence 44, Application US/11106649
 Publication No. US20050181439A1
 GENERAL INFORMATION:
 APPLICANT: Choi et al.

;; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
;; FILE REFERENCE: PB340P2C3D1
;; CURRENT APPLICATION NUMBER: US/11/106,649
;; CURRENT FILING DATE: 2005-04-15
;; PRIOR APPLICATION NUMBER: US 09/765,271
;; PRIOR FILING DATE: 2001-01-22
;; PRIOR APPLICATION NUMBER: US 09/536,784
;; PRIOR FILING DATE: 2000-03-28
;; PRIOR APPLICATION NUMBER: US 08/961,083
;; PRIOR FILING DATE: 1997-10-30
;; PRIOR APPLICATION NUMBER: US 60/029,960
;; PRIOR FILING DATE: 1996-10-31
;; NUMBER OF SEQ ID NOS: 454
;; SOFTWARE: PatentIn Version 3.3
;; SEQ ID NO 44
;; LENGTH: 144
;; TYPE: PRN
;; ORGANISM: Streptococcus pneumoniae
US-11-106-649-44

Query Match 34.0%; Score 282.5; DB 6; Length 144;
Best Local Similarity 44.8%; Pred. No. 1.5e-21;
Matches 65; Conservative 19; Mismatches 56; Indels 5; Gaps 3;

QY 15 GELPVGSPAPAFITLNGDLGVISDQFRGKSVLNIFFSVDTFVPCATSVRTFDER-AAA 73
DB 3 GKQLQVGDKALDPSLTITDLSKKSLLADFDGKKVLSTVPSIDTICSTQTRRPFBEIAGL 62

QY 74 SGATVLCVSKDLFPAGORFGAGESTEN-VMPASAFRDSFGEDYGVTTADGPMGLARAI 132
DB 63 DNTVTVLSMDLFPAGKRCGAGSLDANMLSDYFDHSFGADYALLINEM---HLLARAV 119

QY 133 VVIGADGNVAYTELVEPIAQEPNVE 157
DB 120 FVLDTNTIRYVEYVDINSEPNVE 144

RESULT 14
US-10-335-977-7878
; Sequence 7878, Application US/10335977
; Publication No. US20040052799A1
; GENERAL INFORMATION:
; APPLICANT: DOUGLAS SMITH et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
; RELATING TO HELICOBACTER PYLORI FOR
; DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 10031
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: Windows NT 4.0
; SOFTWARE: UNIX
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/335,977
; FILING DATE: 30-Dec-2002
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/993,002
; FILING DATE: 17-DEC-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandragouras, Amy E.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: GTN-018
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 742-4214
; INFORMATION FOR SEQ ID NO: 7878:

;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 166 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; HYPOTHEICAL: YES
;; ORIGINAL SOURCE:
;; ORGANISM: Helicobacter pylori
;; FEATURE:
;; NAME/KEY: misc feature
;; LOCATION: (8) LOCATION 1...166
;; SEQUENCE DESCRIPTION: SEQ ID NO: 7878:
US-10-335-977-7878

Query Match 30.6%; Score 254; DB 4; Length 166;
Best Local Similarity 35.8%; Pred. No. 1.8e-18;
Matches 59; Conservative 31; Mismatches 73; Indels 2; Gaps 2;

QY 1 MAQITLRGNAINTVGELPVGSPAPAFITLNGDLGVISDQFRGKSVLNIFFSVDTFVPC 60
DB 1 MQRVTFKEETYLEGKALKVGDAPVQLVNGDLQEVNLLKQGRVQVVALPSLTGVC 60

QY 61 ATSVRTFDERAA-ASGATVLCVSKDLFPAGORFGAGESTENVMPASAFR-DSFGEDYGV 118
DB 61 LLOAKAFNEQAGKLPVSYSVISMDFSGQCGICGAGIKDLILDFRYKAFGENYGV 120

QY 119 IADGPMAGLARAIVVIGADGNVAYTELVEPIAQEPNVEAALAL 163
DB 121 LGGSLQGILLARSVFLDQKGVITYKEIVQNLIEEPYVALKL 165

RESULT 15
US-10-335-977-7877
; Sequence 7877, Application US/10335977
; Publication No. US20040052799A1
; GENERAL INFORMATION:
; APPLICANT: DOUGLAS SMITH et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
; RELATING TO HELICOBACTER PYLORI FOR
; DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 10031
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: Windows NT 4.0
; SOFTWARE: UNIX
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/335,977
; FILING DATE: 30-Dec-2002
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/993,002
; FILING DATE: 17-DEC-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandragouras, Amy E.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: GTN-018
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 742-4214
; INFORMATION FOR SEQ ID NO: 7877:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 139 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: YES

```

; ORIGINAL SOURCE:
; ORGANISM: Helicobacter pylori
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...139
; SEQUENCE DESCRIPTION: SEQ ID NO: 7877:
US-10-335-977-7877

```

```

Query Match      26.8%; Score 222.5; DB 4; Length 139;
Best Local Similarity 40.0%; Pred. No. 3e-15;
Matches 52; Conservative 21; Mismatches 46; Indels 11; Gaps 3;

```

```

QY 45 KSVLL-----NIPPSVDTPVCAFSVRTFDRAA-ASGATVLCVSKDLPFAQKRPFG 94
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 9 KSTLEQGVRFQVVSALPSLTGVCVLQAKHFEQAGKLPFSVSISMDELPSQGOICG 68
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 95 AEGTENMPPASAPR-DSFGEDYGTIADGPMAGILARAIVYIGADGNVATTELYPEIAQE 153
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 69 AEGIKDLRIISDFRYKAFGENYGVLLGKSLQGLIARSVFLDDKGVVITYKRIYQNTLBR 128
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 154 PNYEALALAL 163
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 129 PNYEALILKVL 138
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

```

Search completed: April 14, 2006, 18:40:27
 Job time : 61.0026 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using SW model

Run on: April 14, 2006, 18:37:38 ; Search time 7.07037 Seconds
(without alignments)
991.565 Million cell updates/sec

Title: US-10-620-246-8

Perfect score: 831
Sequence: 1 MAQITLRGNAINTVGLPAAV.....LVPEIAGBPNYEALALGA 165

Scoring table: BLASTUM62
Gapop 10.0 , Gapext 0.5

Searched: 217505 seqs, 42489236 residues

Total number of hits satisfying chosen parameters: 217505

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_AA_New:
1: /SIDS5/ptodata/2/pulppaa/US08_NEW_PUB.pep.*
2: /SIDS5/ptodata/2/pulppaa/US06_NEW_PUB.pep.*
3: /SIDS5/ptodata/2/pulppaa/US07_NEW_PUB.pep.*
4: /SIDS5/ptodata/2/pulppaa/PCT_NEW_PUB.pep.*
5: /SIDS5/ptodata/2/pulppaa/US05_NEW_PUB.pep.*
6: /SIDS5/ptodata/2/pulppaa/US10_NEW_PUB.pep.*
7: /SIDS5/ptodata/2/pulppaa/US11_NEW_PUB.pep.*
8: /SIDS5/ptodata/2/pulppaa/US60_NEW_PUB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	467	56.2	172	US-11-079-463-6489	Sequence 6489, Ap
2	466.5	56.1	165	US-10-454-437-172	Sequence 172, App
3	341.5	41.1	123	US-10-454-437-174	Sequence 174, App
4	307	36.9	164	US-10-793-626-252	Sequence 252, App
5	280	33.7	172	US-11-098-686-10841	Sequence 10841, A
6	225	27.1	126	US-11-098-686-81	Sequence 81, Appl
7	215	25.9	164	US-11-074-176-70	Sequence 70, Appl
8	121.5	14.6	257	US-10-642-372A-5	Sequence 5, Appl1
9	121.5	14.6	257	US-10-642-372A-5	Sequence 6, Appl1
10	117.5	14.1	198	US-10-642-372A-28	Sequence 28, Appl
11	117.5	14.1	256	US-10-642-372A-4	Sequence 4, Appl1
12	114.5	13.8	206	US-10-821-234-1038	Sequence 1038, Ap
13	106.5	12.8	193	US-10-527-771-6	Sequence 6, Appl1
14	99	11.9	216	US-11-096-568A-1038	Sequence 1038, Ap
15	99	11.9	226	US-11-096-568A-1037	Sequence 1037, Ap
16	97.5	11.7	318	US-10-506-454-563	Sequence 563, App
17	97	11.7	218	US-11-096-568A-21440	Sequence 21440, A
18	96	11.6	446	US-11-108-172-1121	Sequence 1121, Ap
19	93	10.2	267	US-11-079-463-9249	Sequence 9249, Ap
20	85	10.2	181	US-10-467-657-5556	Sequence 5556, Ap
21	80.5	9.7	271	US-10-821-234-1419	Sequence 1419, Ap
22	80.5	9.7	327	US-11-188-298-8815	Sequence 8815, Ap
23	80	9.6	692	US-11-052-554A-213	Sequence 213, App
24	79.5	9.6	214	US-11-079-463-5569	Sequence 5569, Ap
25	79.5	9.6	233	US-11-096-568A-21082	Sequence 21082, A

26	79.5	9.6	537	US-11-098-686-11361	Sequence 11361, A
27	79	9.5	1461	US-11-052-554A-283	Sequence 283, App
28	75.5	9.1	200	US-11-079-463-7040	Sequence 7040, Ap
29	75.5	9.1	3655	US-11-075-185-5	Sequence 5, Appl1
30	75	9.0	486	US-11-188-298-13811	Sequence 13811, A
31	75	9.0	766	US-11-188-298-17714	Sequence 17714, A
32	73.5	8.8	241	US-10-506-454-206	Sequence 206, App
33	73.5	8.8	337	US-11-188-298-8712	Sequence 8712, Ap
34	73.5	8.8	750	US-11-188-298-15480	Sequence 15480, A
35	72.5	8.7	327	US-11-188-298-13066	Sequence 13066, A
36	72.5	8.7	329	US-11-188-298-15195	Sequence 15195, A
37	72.5	8.7	455	US-11-087-099-6906	Sequence 6906, Ap
38	72	8.7	426	US-11-096-568A-10657	Sequence 10657, A
39	72	8.7	454	US-11-096-568A-10656	Sequence 10656, A
40	72	8.7	919	US-11-188-298-12207	Sequence 12207, A
41	71.5	8.6	167	US-10-467-657-2942	Sequence 2942, Ap
42	71.5	8.6	341	US-11-096-568A-27148	Sequence 27148, A
43	71.5	8.6	830	US-11-096-568A-30527	Sequence 30527, A
44	71.5	8.6	840	US-11-096-568A-30526	Sequence 30526, A
45	71.5	8.6	921	US-11-096-568A-30525	Sequence 30525, A

ALIGNMENTS

```
RESULT 1
US-11-079-463-6489
; Sequence 6489, Application US/11079463
; Publication No. US20050073161A1
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton
; TITLE OR INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO BACTEROIDES FR
; FILE REFERENCE: PATH00-03D1V2
; CURRENT APPLICATION NUMBER: US/11/079,463
; CURRENT FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/128,705
; PRIOR FILING DATE: 1999-04-09
; PRIOR APPLICATION NUMBER: US 09/540,209
; PRIOR FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 10444
; SEQ ID NO 6489
; LENGTH: 172
; TYPE: PRT
; ORGANISM: B.fragilis
US-11-079-463-6489

Query Match      56.2%; Score 467; DB 7; Length 172;
Best Local Similarity 59.4%; Pred. No. 9.4e-40;
Matches 98; Conservative 18; Mismatches 47; Indels 2; Gaps 2;

QY      1 MAQITLRGNAINTVGLPAAVTLTGGLGVISDQFRGKSVLNIFFSDPPVC 60
       7 MATTNKGGPQVKLTIGSFIVGKVAAPDFELVKSLSFALDKDLGKNIVLNIFSLDTGVC 66
QY      61 ATSVRTFDERAAASGAT-VLCVSKDLPFAQKRCFGAEGTENVPAPAFRDS-EGEDYGYT 118
       67 ATSVRKFENKGAAMKOTVLAISKDLPFAQRCCTTEGLENVIFLDFRSDSDSGVR 126
DB      119 IADGPAAGLARAIVIVIGADGNVATYELVPEIAQBPNYEALAL 163
       127 MADGPAAGLARAIVIVIGADGNVATYELVPEIAQBPNYEALAL 171

RESULT 2
US-10-454-437-172
; Sequence 172, Application US/10454437
; Publication No. US20050277115A1
; GENERAL INFORMATION:
; APPLICANT: Pompeius, Markus
; APPLICANT: Kroger, Burkhard
; APPLICANT: Schroder, Hartwig
; APPLICANT: Zeidler, Oskar
```



```

RESULT 5
US-11-098-686-10841
; Sequence 10841, Application US/11098686
; Publication No. US20060024696A1
; GENERAL INFORMATION:
; APPLICANT: Kapur, Vivek and Gebhart, Connie J.
; TITLE OF INVENTION: NUCLEIC ACID AND POLYPEPTIDE SEQUENCES
; TITLE OF INVENTION: FROM LAWSONIA INTRACELLULARIS AND METHODS OF USING
; FILE REFERENCE: 09531-128001
; CURRENT APPLICATION NUMBER: US/11/098,686
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: PCT/US03/31318
; PRIOR FILING DATE: 2003-10-01
; PRIOR APPLICATION NUMBER: US 60/416,395
; PRIOR FILING DATE: 2002-10-04
; NUMBER OF SEQ ID NOS: 11433
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 10841
; LENGTH: 172
; TYPE: PRT
; ORGANISM: Lawsonia intracellularis
US-11-098-686-10841

```

Query March	33.7%	Score 280;	DB 7;	Length 172;
Beet Local Similarity	38.7%;	Pred. No. 6,3e-21;		
Matches	63;	Conservative	26;	Mismatches 68; Indels 6; Gaps 3;
Qy	4	ITLRGNAINTVGELEAVGSPAPAFPLTSGGDISDQFRGKSVLTANTFSPVDTEVCATS	63	
Db	9	ITPLGNPLITLMGNFPISEKAPDPLVLANDSLPRLLSYANNITILSLVPSLDTVCIE	68	
Qy	64	VRTPDERAAAG--ATVLCSKDLFPFAQKRFQGAETENVMPASAFRD--SPGEDYGVTTA	120	
Db	69	TQRFNSBAKLGEEKURLITLISCDLPFAQRMCGAGVAVETLSDRHELSPQYAGIAIK	128	
Qy	121	DGPMAGLARALVIGADGNNAVYTLVPEIAQEPYEEAALAL	163	
Db	129	E---LRLRLAAVFVVDITGMITLYQELIVEMTHPEPYTLIAFEAV	168	

```

RESULT 6
US-11-098-686-81
; Sequence 81, Application US/11098686
; Publication No. US20060024696A1
; GENERAL INFORMATION:
; APPLICANT: Kapur, Vivek and Gebhart, Connie J.
; TITLE OF INVENTION: FROM LAMSONIA INTRACELLULARIS AND METHODS OF USING
; TITLE OF INVENTION: FROM LAMSONIA INTRACELLULARIS AND METHODS OF USING
; FILE REFERENCE: 09531-128001
; CURRENT APPLICATION NUMBER: US/11/098,686
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: PCT/US03/31318
; PRIOR FILING DATE: 2003-10-01
; PRIOR APPLICATION NUMBER: US 60/416,395
; PRIOR FILING DATE: 2002-10-04
; NUMBER OF SEQ ID NOS: 11433
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 81
; LENGTH: 126
; TYPE: PRT
; ORGANISM: Lamsonia intracellularis
US-11-098-686-81

```

Query Match	27.1%	Score 225	DB 7	length 126
Best Local Similarly	42.9%	Pred. No. 1.5e-15		
Matches 51; Conservative	18	Mismatches 44	Indels 6	Gaps 3

[illegible]

```

Db                               71 SDHRELSTFGVATGAIKE---LRLLARAVFVVDNCGMTTYGELVPEMTHETPTALFFEA 126

RESULT 7
US-11-074-176-70
; Sequence 70, Application US/11074176
; Publication No. US20050250135A1
; GENERAL INFORMATION:
; APPLICANT: Klaenhammer, Todd R.
; APPLICANT: Russell, William M.
; APPLICANT: Altermann, Eric
; APPLICANT: Mcauliffe, Olivia
; APPLICANT: Perill, Andrea Azcarate
; TITLE OF INVENTION: Nucleic Acid Sequences Encoding
; TITLE OF INVENTION: Stress-Related Proteins and Uses Therefore
; FILE REFERENCE: 5051-694
; CURRENT APPLICATION NUMBER: US/11/074,176
; CURRENT FILING DATE: 2005-03-07
; PRIOR APPLICATION NUMBER: 60/551,161
; PRIOR FILING DATE: 2004-03-08
; NUMBER OF SEQ ID NOS: 381
; SOFTWARE: FastSeq for windows Version 4.0
; SEQ ID NO 70
; LENGTH: 164
; TYPE: PRT
; ORGANISM: Lactobacillus acidophilus
US-11-074-176-70

```

[illegible]

```

RESULT 8
US-10-642-272A-5
; Sequence 5, Application US/10642272A
; Publication No. US20050277606A1
; GENERAL INFORMATION:
; APPLICANT: Hattori, Fumiyuki
; APPLICANT: Sugimura, Keiichi
; APPLICANT: Furuya, Mayumi
; TITLE OF INVENTION: Therapeutic Methods and Agents for Diseases Associated with
; TITLE OF INVENTION: Decreased Expression of AOP-1 Gene or AOP-1
; FILE REFERENCE: 58777.000012
; CURRENT APPLICATION NUMBER: US/10/642,272A
; CURRENT FILING DATE: 2003-08-18
; PRIOR APPLICATION NUMBER: PCT/JP02/01358
; PRIOR FILING DATE: 2001-02-18
; PRIOR APPLICATION NUMBER: JP 41003/2001
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 5
; LENGTH: 257
; TYPE: PR1
; ORGANISM: Rattus norvegicus
; US-10-642-272A-5

```

Query Match	Score	DB 6;	Length
14.6%;	121.5;		257;
Best Local Similarity	26.8%;	Pred. No. 9.6e-05;	

APPLICANT: Labat, Ivan
APPLICANT: Seache-Crafn, Birgit
APPLICANT: Andarmat, Susan
APPLICANT: Tang, Y. Tom
TITLE OF INVENTION: Methods for diagnosis and treatment of Preeclampsia
FILE REFERENCE: 821A
CURRENT APPLICATION NUMBER: US/10/821.234
PRIOR FILING DATE: 2004-04-07
PRIOR APPLICATION NUMBER: US 60/462,047
PRIOR FILING DATE: 2003-04-07
NUMBER OF SEQ ID NOS: 1704
SOFTWARE: pt_seq_genes Version 1.0
SEQ ID NO 1038
LENGTH: 206
TYPE: PRT
ORGANISM: Homo sapiens
US-10-821-234-1038

Query Match 13.8%; Score 114.5; DB 6; Length 206;
Best Local Similarity 29.9%; Pred. No. 0.0037;
Matches 43; Conservative 12; Mismatches 66; Indels 23; Gaps 5;

QY 8 GNAINTVGLPVAWSBAPATLTG---GDLGVISDQFRGKSVLTNIIPSVDTVCATSV 64
DB 12 GNA-----RIGKPAIDFKATVVDGAFKRVKLSYKGYVLPFYPLDFTVCPTETI 63
QY 65 RTFDERA---AASGATVLCVSKDLPFAQ-----KRFCAEGTENMPASAFRDSFGADY 115
DB 64 IASNRABEDPRKACGCVLGSVDSQFTHLAMINTPRKGGIGLPIATLADVTARLSBDY 123
QY 116 GVTIADGPMAGLLARAIVIGADG 139
DB 124 GVLKTD---EGIAVRGLFIIDKG 144

RESULT 13
US-10-527-771-6
Sequence 6, Application US/10527771
Publication No. US20050271683A1
GENERAL INFORMATION:
APPLICANT: University Gent
TITLE OF INVENTION: Ostertagia vaccine
FILE REFERENCE: 2002-015
CURRENT APPLICATION NUMBER: US/10/527,771
CURRENT FILING DATE: 2005-03-11
PRIOR APPLICATION NUMBER: US 10/243,319
PRIOR FILING DATE: 2002-09-13
NUMBER OF SEQ ID NOS: 27
SOFTWARE: PatentIn version 3.2
SEQ ID NO 6
LENGTH: 193
TYPE: PRT
ORGANISM: Ostertagia ostertagi
FEATURES:
NAME/KEY: misc.feature
LOCATION: (72)-(72)
OTHER INFORMATION: The xaa at location 72 stands for Asn, or Thr.
US-10-527-771-6

Query Match 12.8%; Score 106.5; DB 6; Length 193;
Best Local Similarity 27.3%; Pred. No. 0.0021;
Matches 36; Conservative 16; Mismatches 65; Indels 15; Gaps 4;

QY 20 VGSAPAPF---TLTGGDLGVISDQFRGKSVLTNIIPSVDTVCATSVRTFDERA---AA 73
DB 3 IGPAPDPATKAVYNGDFIDVKSVDYKXTVLFYPLDFTVCPTETIIPSDRVEFPK 62
QY 74 SGATVLCVSKDLPFAQ-----KRFCAEGTENMPASAFRDSFGADYGTIADGPMAGL 127
DB 63 IDAAVLAAGSDVSFSLAMINTPRKMGGLDPMIPVATDNRHQIAKDYGLAKED---EGI 119
QY 128 LARAIVIGADG 139

DB 120 AYRGLFIIDPKG 131

RESULT 14
US-11-096-568A-1038
Sequence 1038, Application US/11096568A
Publication No. US20060048240A1
GENERAL INFORMATION:
APPLICANT: Alexandrov, Nikolai et al.
TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
FILE REFERENCE: 2750-1592PUS2
CURRENT APPLICATION NUMBER: US/11/096,568A
CURRENT FILING DATE: 2005-04-01
NUMBER OF SEQ ID NOS: 34471
SEQ ID NO 1038
LENGTH: 216
TYPE: PRT
ORGANISM: Zea mays subsp. mays
FEATURES:
NAME/KEY: misc.feature
LOCATION: (1)-(216)
OTHER INFORMATION: Cereals Seq. ID no. 13599403
US-11-096-568A-1038

Query Match 11.9%; Score 99; DB 7; Length 216;
Best Local Similarity 28.1%; Pred. No. 0.014;
Matches 36; Conservative 25; Mismatches 53; Indels 14; Gaps 5;

QY 21 GSPAPAFTLTGGDLGVISDQFRGKSVLTNIIPSVDTVC---ATSVRTFDERAASGAT 77
DB 72 GGSAPDFTLKDQNGKVPVSLKRYKGPVYVYFPADETPGCTKQACAFRDSYEKKKAGAB 131
QY 78 VLCVSKDLPFAQRFCAGETENMPASAFRDS---FGEDYGTIADGPMAGLL-ARAIV 133
DB 132 VIGISGDSASHKAF---ASKTKLFTLLSDGNRVKDMGV---RDLFGALPGHOTY 184
QY 134 VTGADGNV 141
DB 185 VLDRNGV 192

RESULT 15
US-11-096-568A-1037
Sequence 1037, Application US/11096568A
Publication No. US20060048240A1
GENERAL INFORMATION:
APPLICANT: Alexandrov, Nikolai et al.
TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
FILE REFERENCE: 2750-1592PUS2
CURRENT APPLICATION NUMBER: US/11/096,568A
CURRENT FILING DATE: 2005-04-01
NUMBER OF SEQ ID NOS: 34471
SEQ ID NO 1037
LENGTH: 226
TYPE: PRT
ORGANISM: Zea mays subsp. mays
FEATURES:
NAME/KEY: misc.feature
LOCATION: (1)-(226)
OTHER INFORMATION: Cereals Seq. ID no. 13599402
US-11-096-568A-1037

Query Match 11.9%; Score 99; DB 7; Length 226;
Best Local Similarity 28.1%; Pred. No. 0.015;
Matches 36; Conservative 25; Mismatches 53; Indels 14; Gaps 5;

QY 21 GSPAPAFTLTGGDLGVISDQFRGKSVLTNIIPSVDTVC---ATSVRTFDERAASGAT 77
DB 82 GGSAPDFTLKDQNGKVPVSLKRYKGPVYVYFPADETPGCTKQACAFRDSYEKKKAGAB 141
QY 78 VLCVSKDLPFAQRFCAGETENMPASAFRDS---FGEDYGTIADGPMAGLL-ARAIV 133

Db	142	VIGISGDBSASHKAF----	ASKYKLPYTLSDGNNRVKDWGV--	FGDLFGALPGRQTY	194
Qy	134	VIGADGNV	141		
Db	195	VLDKNGV	202		

Search completed: April 14, 2006, 18:42:07
Job time : 8.07037 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 14, 2006, 17:03:40 ; Search time 70.5178 Seconds
(without alignments)
1133.996 Million cell updates/sec

Title: US-10-620-246-12

Perfect score: 961
Sequence: 1 MADCDSTVNSPLATATATLH.....ATDGNDRPTDPVVISITIS 182

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_21.*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*
9: geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	961	100.0	182	2	AAW72889 Mycobacte
2	961	100.0	182	2	AAW72889 Mycobacte
3	716	74.5	187	4	AAW72889 Mycobacte
4	607	63.2	182	4	AAW72889 Mycobacte
5	494.5	51.5	182	6	ABM41697 Propionib
6	465.5	48.4	182	5	ABM41697 Propionib
7	460.5	47.9	188	9	ABM41697 Propionib
8	460.5	47.9	205	9	ABM41697 Propionib
9	431.5	44.9	180	2	AAW4191 Cyclophil
10	407.5	42.4	466	6	ABU01998 S. pneumo
11	407.5	42.4	466	6	ABU01998 S. pneumo
12	407.5	42.4	466	6	ABU01998 S. pneumo
13	407.5	42.4	466	6	ABU01998 S. pneumo
14	403.5	42.0	466	8	ADK46830 Streptoco
15	403.5	42.0	472	8	ADK46830 Streptoco
16	403.5	42.0	472	8	ADK46830 Streptoco
17	394.5	41.1	196	5	ABM41697 Propionib
18	393	40.9	466	6	ABU01998 S. pneumo
19	392	40.8	470	5	ABM41697 Propionib
20	388	40.4	452	8	ADK46830 Streptoco
21	388	40.4	452	8	ADK46830 Streptoco
22	382	39.8	165	3	AAW72889 Mycobacte
23	382	39.8	165	3	AAW72889 Mycobacte
24	381	39.6	639	8	ADT58020 plant pol

25	374	38.9	160	3	AAW72889 Mycobacte
26	373	38.8	193	3	AAW72889 Mycobacte
27	370	38.5	160	3	AAW72889 Mycobacte
28	359.5	37.4	420	4	AAW72889 Mycobacte
29	359.5	37.4	420	8	ADK46830 Streptoco
30	359.5	37.4	420	8	ADK46830 Streptoco
31	359.5	37.4	420	8	ADK46830 Streptoco
32	359.5	37.4	420	8	ADK46830 Streptoco
33	359.5	37.4	420	8	ADK46830 Streptoco
34	359.5	37.4	420	8	ADK46830 Streptoco
35	359.5	37.4	420	8	ADK46830 Streptoco
36	359.5	37.4	420	8	ADK46830 Streptoco
37	359.5	37.4	420	8	ADK46830 Streptoco
38	359.5	37.4	420	8	ADK46830 Streptoco
39	359.5	37.4	420	8	ADK46830 Streptoco
40	359.5	37.4	420	8	ADK46830 Streptoco
41	358.5	37.3	199	6	ADK46830 Streptoco
42	358.5	37.3	202	6	ADK46830 Streptoco
43	358.5	37.3	207	6	ADK46830 Streptoco
44	352	36.6	166	4	AAW72889 Mycobacte
45	352	36.6	166	6	ABO17619 Novel hum

ALIGNMENTS

RESULT 1	AAW72889	standard; protein; 182 AA.
XX	AAW72889;	
XX	21-JUN-1999 (first entry)	
XX	Mycobacterium tuberculosis antigen CFP22.	
XX	Mycobacterium tuberculosis; antigen; vaccine; immunological; immunogen; infection.	
XX	Mycobacterium tuberculosis.	
XX	MO9844119-A1.	
XX	08-OCT-1998.	
XX	01-APR-1998; 98WO-DK000132.	
XX	02-APR-1997; 97DK-00000376.	
XX	18-APR-1997; 97US-0044624P.	
XX	10-NOV-1997; 97DK-00001277.	
XX	05-JUN-1998; 98US-0070488P.	
XX	(STAT-) STATENS SERUM INST.	
XX	Andersen P, Nielsen R, Rosenkrands I, Weldingh K, Rasmussen PB, Oettinger T, Florio W;	
XX	WPI, 1998-542705/46.	
XX	N-PSDB; AAW63920.	
XX	New isolated mycobacteria polypeptides and nucleic acids - used for developing products for the diagnosis of or vaccination against mycobacterial infections, particularly tuberculosis.	
XX	Claim 1; Page 134-135; 163pp; English.	
XX	The present sequence represents a Mycobacterium tuberculosis protein. Products from the present invention, which describes protein fragments and nucleic acid fragments derived from M. tuberculosis, can be used in the detection of and prevention of mycobacterial infections. In particular, the proteins and nucleic acids can be used for the diagnosis of or vaccination against tuberculosis caused by M. tuberculosis, M. africanum or M. bovis	

```

XX      SQ      Sequence 182 AA;
Query Match      100.0%; Score 961; DB 2; Length 182;
Best Local Similarity 100.0%; Pred. No. 3,2e-93;
Matches 182; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MADCSVTNSPLATATATLHTNRGDIKIALFGNHAKPTVANFVGLAQGTQDYSTQNASGG 60
DB      1 MADCSVTNSPLATATATLHTNRGDIKIALFGNHAKPTVANFVGLAQGTQDYSTQNASGG 60
QY      61 PSGPFYDGAVFHRVIOGFMIQGGDPGTGGRGPGYKFADEBFHPELOFDPKPYLLAMNAGP 120
DB      61 PSGPFYDGAVFHRVIOGFMIQGGDPGTGGRGPGYKFADEBFHPELOFDPKPYLLAMNAGP 120
QY      121 GTNGSQFITYGKTPHLNRRHTIFGEVIDAESQRVVEAISKTATDGNDRPTDPVIESIT 180
DB      121 GTNGSQFITYGKTPHLNRRHTIFGEVIDAESQRVVEAISKTATDGNDRPTDPVIESIT 180
QY      181 IS 182
DB      181 IS 182

RESULT 2
AA21906 standard; protein; 182 AA.
XX      AA21906;
AC      06-SBP-1999 (first entry)
DT
DE      Amino acid sequence of antigen CFP22.
XX
XX      Immunogenic; Mycobacterium tuberculosis; immune response; infection;
KM      tuberculosis; fusion polypeptide; T-cell epitope; BSAT-6; MPT59; TB;
KM      pharmaceutical; vaccination; M. africanum; M. bovis; CFP7A; CFP30A;
KM      CFP7B; CFP19; CFP27; CFP30A; RDI-ORF; CFP10A; CFP16; CFP19; CFP23;
KM      CFP25A; CFP30B; CFP7B.
XX
XX      Mycobacterium tuberculosis.
OS
XX
XX      WO924577-A1.
PN
XX      20-MAY-1999.
PD
XX
XX      08-OCT-1998; 98WO-DK000438.
PF
XX      10-NOV-1997; 97DK-00001277.
PR      05-JAN-1998; 98US-0070488P.
PR      01-APR-1998; 98WO-DK000132.
XX
XX      (STAT-) STATENS SERUM INST.
PA
XX      Andersen P, Skjot R;
PI
XX      WPI; 1999-347282/29.
DR      N-PSDB; AAX81006.
XX
XX      New immunogenic fragment of Mycobacterium tuberculosis.
PT
XX
XX      Example 3; Page 60; 265pp; English.
PS
XX      The invention describes a substantially pure immunogenic polypeptide
CC      fragment (1) from Mycobacterium tuberculosis that is able to evoke a
CC      protective immune response against infections by mycobacteria belonging
CC      to the tuberculosis complex. The invention provides a (1) fusion
CC      polypeptide comprising at least one polypeptide fragment (1) and at least
CC      one fusion partner; (2) a fusion polypeptide fragment comprising a T-cell
CC      epitope from M. tuberculosis protein BSAT-6, or MPT59 and a second
CC      different amino acid sequence from M. tuberculosis, and/or including a
CC      sequence which protects the first amino acid sequence from in vivo
CC      degradation or post-translational processing; (3) a nucleic acid fragment

```

```

CC      that encodes the above polypeptides. The polypeptides and nucleic acid
CC      are useful as pharmaceuticals, for diagnosis of and as antigens for
CC      vaccination against TB caused by Mycobacterium tuberculosis, africanum or
CC      bovis. The polypeptides are also useful for diagnosing ongoing or
CC      previous sensitization in an animal with bacteria belonging to the
CC      tuberculosis complex. The invention also describes the use of CFP7A or
CC      CFP30A or a T-cell epitope of for the induction of a strong immune
CC      response in a mammal; use of CFP7B, CFP19 or MPT59-BSAT6 or a T-cell
CC      epitope of for diagnosis of TB in a mammal by performing a DTH type skin
CC      test; use of CFP27, CFP30A, RDI-ORF2, RDI-ORF5, MPT59-BSAT6,
CC      BSAT6-MPT59, CFP10A, CFP16, CFP19, CFP23, CFP25A, CFP30B, CFP7B or a T-
CC      cell epitope of for the preparation of an immunological composition; and
CC      for the preparation of a subunit vaccine.
XX
XX      SQ      Sequence 182 AA;
Query Match      100.0%; Score 961; DB 2; Length 182;
Best Local Similarity 100.0%; Pred. No. 3,2e-93;
Matches 182; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MADCSVTNSPLATATATLHTNRGDIKIALFGNHAKPTVANFVGLAQGTQDYSTQNASGG 60
DB      1 MADCSVTNSPLATATATLHTNRGDIKIALFGNHAKPTVANFVGLAQGTQDYSTQNASGG 60
QY      61 PSGPFYDGAVFHRVIOGFMIQGGDPGTGGRGPGYKFADEBFHPELOFDPKPYLLAMNAGP 120
DB      61 PSGPFYDGAVFHRVIOGFMIQGGDPGTGGRGPGYKFADEBFHPELOFDPKPYLLAMNAGP 120
QY      121 GTNGSQFITYGKTPHLNRRHTIFGEVIDAESQRVVEAISKTATDGNDRPTDPVIESIT 180
DB      121 GTNGSQFITYGKTPHLNRRHTIFGEVIDAESQRVVEAISKTATDGNDRPTDPVIESIT 180
QY      181 IS 182
DB      181 IS 182

RESULT 3
AA689785 standard; protein; 187 AA.
XX      AA689785;
AC      26-SBP-2001 (first entry)
DT
DE      C glutamicum protein fragment SEQ ID NO: 3539.
XX
XX      Corynebacterium; amino acid synthesis; vitamin; saccharide;
KM      organic acid synthesis.
XX
XX      Corynebacterium glutamicum.
OS
XX
XX      BP1108790-A2.
PN
XX      20-JUN-2001.
PD
XX
XX      18-DEC-2000; 2000EP-00127688.
PF
XX      16-DEC-1999; 99JP-00377484.
PR      07-APR-2000; 2000JP-00159162.
PR      03-AUG-2000; 2000JP-00280988.
XX
XX      (KYOW ) KYOWA HAKKO KOGYO KK.
PA
XX      Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
PI      Tateishi N, Senoh A, Ikeda M, Ozaki A;
XX
XX      WPI; 2001-376931/40.
DR      N-PSDB; AA65004.
XX
XX      Novel polynucleotides derived from Corynebacterium bacteria, for identifying
PT      mutation point of a gene, measuring expression of a gene, analyzing
PT      expression profile or pattern of a gene and identifying homologous gene.

```

XX Claim 17; SEQ ID NO 3539; 246bp + Sequence Listing; English.
 PS
 XX
 CC The present invention provides a number of nucleotide and protein
 CC sequences from the *Corynebacterium* *Corynebacterium* glutamicum. These
 CC are useful for identifying the mutation point of a gene derived from a
 CC mutant of *Corynebacterium* *Corynebacterium*, measuring expression amount and analyzing
 CC the expression profile or expression pattern of a gene derived from
 CC *Corynebacterium* *Corynebacterium*, and identifying a homolog of a gene derived from
 CC *Corynebacterium* *Corynebacterium*. *Corynebacterium* bacteria are useful for producing amino
 CC acids, nucleic acids, vitamins, saccharides and organic acids,
 CC particularly L-lysine. The present sequence is a protein described in the
 CC exemplification of the invention. Note: The sequence data for this patent
 CC did not form part of the printed specification, but was obtained in
 CC electronic format directly from the European Patent Office
 XX
 SQ Sequence 187 AA;
 Query Match 74.5%; Score 716; DB 4; Length 187;
 Best Local Similarity 79.9%; Pred. No. 2.9e-67;
 Matches 135; Conservative 11; Mismatches 23; Indels 0; Gaps 0;
 QY 14 TATATLHTNRGDIKIALFGNHAAPKTVANFVGLAGTGYSTONASGSGSPFYDGA VPHR 73
 DB 18 TATATLHTNRGDIKIALFGNHAAPKTVANFVGLAGTGYSTONASGSGSPFYDGA VPHR 77
 QY 74 VIQGFMIQGGDPPTGGRGPGYKXPADEFHPELOPKPYLLMANANGPPTNGSQFFITVTK 133
 DB 78 VIQGFMIQGGDPPTGGRGPGYKXPADEFHPELOPKPYLLMANANGPPTNGSQFFITVTP 137
 QY 134 TPHLNRHTTIFGEVVIDAESQKRVFAISKTATDGNDRPTDPVVISITIS 182
 DB 138 TPHLNRHTTIFGEVVIDAESQKRVFAISKTATDGNDRPTDPVVISITIS 186
 RESULT 4
 AAU45178 ID AAU45178 standard; protein; 182 AA.
 AC AAU45178;
 DT 27-FEB-2002 (first entry)
 XX
 DE Propionibacterium acnes immunogenic protein #6074.
 XX
 KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
 KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
 KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
 KW dermatological; osteopathic; neuroprotectant.
 XX
 OS Propionibacterium acnes.
 XX
 PN WO200181581-A2.
 XX
 PD 01-NOV-2001.
 XX
 PF 20-APR-2001; 2001WO-US012865.
 XX
 PR 21-APR-2000; 2000US-0199047P.
 PR 02-JUN-2000; 2000US-0208841P.
 PR 07-JUL-2000; 2000US-0216747P.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 PI Skeiky YAM, Persing DH, Mitcham JL, Wang SS, Bhactia A;
 PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
 DR WPI; 2001-616774/71.
 DR N-PSDB; AAS59525.
 XX
 PT Propionibacterium acnes polypeptides and nucleic acids useful for
 PT vaccinating against and diagnosing infections, especially useful for
 PT treating acne vulgaris.

XX Claim 6; SEQ ID NO 6373; 1069bp; English.
 PS
 XX
 CC Sequences AAU39105-AAU68017 represent *Propionibacterium* *Propionibacterium* acnes immunogenic
 CC polypeptides. The proteins and their associated DNA sequences are used in
 CC the treatment, prevention and diagnosis of medical conditions caused by
 CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
 CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
 CC P. acnes is also involved in infections of bone, joints and the central
 CC nervous system, however it is particularly involved in the inflammatory
 CC lesions associated with acne vulgaris. A method for detecting the
 CC presence or absence of P. acnes in a patient comprises contacting a
 CC sample with a binding agent that binds to the proteins of the invention
 CC and determining the amount of bound protein in the sample. The
 CC polypeptides may be used as antigens in the production of antibodies
 CC specific for P. acnes proteins. These antibodies can be used to
 CC downregulate expression and activity of P. acnes polypeptides and
 CC therefore treat P. acnes infections. The antibodies may also be used as
 CC diagnostic agents for determining P. acnes presence, for example, by
 CC enzyme linked immunosorbent assay (ELISA). Note: The sequence data for
 CC this patent did not form part of the printed specification, but was
 CC obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 182 AA;
 Query Match 63.2%; Score 607; DB 4; Length 182;
 Best Local Similarity 68.5%; Pred. No. 9.9e-56;
 Matches 115; Conservative 16; Mismatches 37; Indels 0; Gaps 0;
 QY 15 ATATLHTNRGDIKIALFGNHAAPKTVANFVGLAGTGYSTONASGSGSPFYDGA VPHR 74
 DB 15 STATLRNHDGDIKIALFGNHAAPKTVANFVGLAGTGYSTONASGSGSPFYDGA VPHR 74
 QY 75 ICGFMIGGDPPTGGRGPGYKXPADEFHPELOPKPYLLMANANGPPTNGSQFFITVTK 134
 DB 75 ICGFMIGGDPPTGGRGPGYKXPADEFHPELOPKPYLLMANANGPPTNGSQFFITVAP 134
 QY 135 PHLNRHTTIFGEVVIDAESQKRVFAISKTATDGNDRPTDPVVISITIS 182
 DB 135 PHLNRHTTIFGEVVIDAESQKRVFAISKTATDGNDRPTDPVVISITIS 182
 RESULT 5
 ABM41697 ID ABM41697 standard; protein; 182 AA.
 AC ABM41697;
 DT 20-OCT-2003 (first entry)
 XX
 DE Propionibacterium acnes immunogenic polypeptide #6373.
 XX
 KW Acne vulgaris; anti-seborrheic; dermatological; antibacterial;
 KW immunostimulant; immune response; vaccine; immunogenic.
 XX
 OS Propionibacterium acnes.
 XX
 PN WO2003033515-A1.
 XX
 PD 24-APR-2003.
 XX
 PF 11-OCT-2002; 2002WO-US032727.
 XX
 PR 15-OCT-2001; 2001US-00978825.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 PI Mitcham JL, Skeiky YAM, Persing DH, Bhactia A, Maisonneuve JL;
 PI Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D;
 PI Barch B, Valliave-Douglas J;
 DR WPI; 2003-381789/36.

[illegible]

```

PF 10-JUL-2001; 2001US-00902540.
XX
PR 10-JUL-2000; 2000US-0217883P.
XX
PA (MONS ) MONSANTO TECHNOLOGY LLC.
XX
PI Goldman BS, Hinkle GJ, Slater SC, Wiegand RC;
XX
DR WPI; 2005-028716/03.
XX
PT New substantially purified Myxococcus xanthus nucleic acid molecule
PT encoding a nitrite reductase, useful for determining gene expression,
PT identifying mutations in a gene of interest, and for constructing
PT mutations in a gene of interest.
XX
PS Example 2; SEQ ID NO 13998; 25bp; English.
XX
CC The invention relates to a substantially purified nucleic acid molecule
CC encoding a nitrite reductase of SEQ ID NO. 11926. Further disclosed is a
CC recombinant DNA construct for expression of a nitrite reductase gene in a
CC plant cell, and a plant cell comprising the recombinant DNA construct.
CC The nucleic acid is useful for determining gene expression, identifying
CC mutations in a gene of interest, and for constructing mutations in a gene
CC of interest. Sequences given in records for SEQ IDs 9692-16825 represent
CC a group of 7134 Myxococcus xanthus proteins and peptides. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from USPTO
XX
SQ Sequence 182 AA;
XX
XX
Query Match 51.5%; Score 494.5; DB 9; Length 182;
Best Local Similarity 60.5%; Pred. No. 8.2e-44;
Matches 101; Conservative 12; Mismatches 51; Indels 3; Gaps 2
XX
QY 17 ATLTNRGDIKIALFGNHAPKTVANFVGLAOSTKDYSTONASGCBPG--PFYDGAVFHRYI 75
DB 9 ATLETNGAIVYRLFSKDAKPTVANKFGLATGEKAMTDPKTGQRVEKRLYDGVIFHRYI 68
QY 76 QGFPMIQGDPPTGTGGRGSPGYKPADBFHELOFDKPYLLAMANAGFQTNGSOFFITVGKTP 135
DB 69 PGFPMIQGDDPTGTGRGDPGYRFEDEFQSGRPFDDKGLLAMANAGGQTNGSOFFITSTPD 128
QY 136 HLNRRHTTFGRVIDAESGRVVEALSKVATDNGDRPTDVPVIESITIS 182
DB 129 YLNNRHTTFGRVSVGYD--VVEKISVNGVRDPRDKLEBPVVIQIKLMS 173
XX
RESULT 7
ABP66140
ID ABP66140 standard; protein; 182 AA.
XX
AC ABP66140;
XX
DT 19-NOV-2002 (first entry)
XX
DE Bifidobacterium longum NCC2705 ORF amino acid sequence SEQ ID NO.884.
XX
KW Bifidobacterium longum NCC2705; Bifidobacterium; bacterial;
XX
KW antidiarrhetic; antibacterial; inhibitor of Salmonella; detection;
XX
KW identification; lactic acid bacterium; diarrhoea; pathogenic bacteria;
XX
KW rotavirus; food composition; pharmaceutical composition.
XX
OS Bifidobacterium longum.
XX
PN EP1227152-A1.
XX
PD 31-JUL-2002.
XX
PF 30-JAN-2001; 2001EP-00102050.
XX
PR 30-JAN-2001; 2001EP-00102050.
XX
PA (NEST ) SOC PROD NESTLE SA.

```


XX 21-NOV-2003; 2003FR-00013687.
 PR (INSP) INST PASTEUR.
 XX (INRM) INSERM INST NAT SANTE & RECH MEDICALE.
 PA (UYLY-) UNIT LYON I BERNARD CLAUDE.
 PA (CNRS) CNRS CENT NAT RECH SCI.
 XX Buchteser C, Tichit M, Etienne J, Ma L, Cazalet C, Glaeser P,
 PI Ruenlok C, Bouchier C, Zidane N, Magnier A, Kunat F, Vandenesch F,
 PI Jarraud S;
 DR WPI; 2005-388305/40.
 XX
 PT New genome of *Legionella pneumophila* Paris strain and derived
 PT polypeptides, useful for detection or identification of the strain and
 PT for treatment and prevention of infections.
 XX
 PS Claim 3; SEQ ID NO 2657; 660pp; English.
 XX
 CC The invention relates to an isolated or purified nucleotide sequences (I)
 CC from *Legionella pneumophila* Paris strain. (II), and their related
 CC sequences or fragments, are useful as primers and probes for detection
 CC and amplification, including differentiation between the Paris and
 CC Philadelphia serains of *Legionella pneumophila* and to prepare recombinant
 CC (hybrid) polypeptides (II). (II) are also useful for preparation of
 CC specific antibodies (Ab), also used for detection/identification of
 CC *Legionella*, and some (II), specifically those involved in synthesis of
 CC surface proteins, are targets for identification of inhibitors. (II), or
 CC vectors that contain (I), are useful as vaccines and immunogenic
 CC compositions, for treatment and prevention of infections by *L.*
 CC *pneumophila*. The present sequence represents the amino acid sequence of a
 CC *L. pneumophila* protein.
 XX
 SQ Sequence 205 AA;
 Query Match 47.9%; Score 460.5; DB 9; Length 205;
 Best Local Similarity 56.6%; Pred. No. 3.8e-40;
 Matches 94; Conservative 17; Mismatches 52; Indels 3; Gaps 2;
 QY 16 TATLHNRGDIKALFGNNAKTVANFVGLAOGTKDYSTONAGSGPGSPYDGAVPFRV 75
 DB 41 TALIKTSEGNITCELPFKENPNTVANFVGLATGKEDVKTSBMVRKPPYNGILIFRVI 100
 QY 76 OGFMIOGDPPTGTGRCGPGYKFADEFHPELODKPYLLAMANAQPGTNGSPFTVGTPT 135
 DB 101 AGFMIOGDPPLGNGTGTGPGTTPDNE-NTNASTFKPGVLLAMANAQPGTNGSPFTTVAPTP 159
 QY 136 HLMRRHTTIGEVIDAESQRYVEAISKATDGNDRPTDPPVVISITTI 181
 DB 160 ELQGNYNVFGQVI--SGQEVVDKISKMPDPPDKPKIPVIVIENTITI 203
 RESULT 10
 AAW4191
 ID AAW4191 standard; protein; 180 AA.
 XX
 AC AAW4191;
 XX
 DT 17-OCT-2003 (revised)
 DT 08-JUN-1998 (first entry)
 XX
 DE Cyclophilin type PPIase.
 DE
 XX Cyclophilin PPIase; halophilic; archaeobacterium; immunosuppressant;
 KM cyclosporin A.
 KM
 XX Halobacterium salinarum.
 OS
 XX JP09313184-A.
 XX
 PN 09-DEC-1997.
 PD
 XX

PF 28-MAY-1996; 96JP-00133353.
 XX
 XX 28-MAY-1996; 96JP-00133353.
 XX
 PA (KAIY-) KAIYO BIOTECHNOLOGY KENKYUSHO KK.
 XX
 XX WPI; 1998-080075/08.
 DR N-PSDB; AAV12290.
 DR
 XX New cyclophilin type PPIase gene - purified from halophilic
 PT archaeobacterium.
 PT
 XX
 PS Example 5; Fig 1; 6pp; Japanese.
 PS
 XX The present sequence represents a cyclophilin type PPIase gene from a
 CC halophilic archaeobacterium, Halobacterium cutirubrum. The cyclophilin
 CC type PPIase may be combined with an immunosuppressant cyclosporin A.
 CC (updated on 17-OCT-2003 to standardise OS field)
 XX
 SQ Sequence 180 AA;
 Query Match 44.9%; Score 431.5; DB 2; Length 180;
 Best Local Similarity 50.0%; Pred. No. 3.8e-37;
 Matches 88; Conservative 20; Mismatches 55; Indels 13; Gaps 3;
 QY 17 ATLHNRGDIKALFGNNAKTVANFVGLAOGTKDYSTONAGSGPGSP-PP 65
 DB 5 ATVHTSEGFDELYDERAPRTVENFLNLRHPPADPADPADPTVWEDPESGEIRGDSL 64
 QY 66 YDGAVPFRVIOGFMIOGDPPTGTGRCGPGYKFADEFHPELODKPYLLAMANAQPGTNGS 125
 DB 65 YAVSVSHRIIEGMIQSGDPTGTGRCGPGYKFADEFHPELODKPYLLAMANAQPGTNGS 124
 QY 126 QFRTVGTGKTPHLNRHRTTIGEVIDAESQRYVEAISKATDGNDRPTDPPVVISITTI 181
 DB 125 QFRTTIDAGPHLDGRNRAVGAATD--GMDVETIGVETIDANDAPASEITIDRVEI 178
 RESULT 11
 ID ABU01998
 ABU01998 standard; protein; 466 AA.
 XX
 AC ABU01998;
 XX
 DT 23-OCT-2003 (revised)
 DT 11-FEB-2003 (first entry)
 XX
 DE S. pneumoniae type 4 strain protein from coding region #1575.
 DE
 XX Bacterial meningitis; pneumonia; sepsis; otitis media; ear infection;
 KM antiinflammatory; antibacterial; immunostimulant; auditory; respiratory;
 KM gene therapy; vaccine.
 XX
 OS Streptococcus pneumoniae; type 4 strain.
 OS
 XX WO200277021-A2.
 PN
 XX 03-OCT-2002.
 PD
 XX 27-MAR-2002; 2002WO-1B002163.
 XX
 XX 27-MAR-2001; 2001GB-00007658.
 PR
 XX (CHIR-) CHIRON SPA.
 PA (GENO-) INST GENOMIC RES.
 PA
 XX Masignani V, Tettelin H, Frazer C;
 PT
 XX WPI; 2003-040579/03.
 DR N-PSDB; ABX07287.
 DR
 XX New proteins and nucleic acid molecules from *Streptococcus pneumoniae*,
 PT useful as medicaments for treating or preventing a disease or infection

PT due to streptococcus bacteria, such as pneumonia, sepsis, otitis media or ear infection.

PS Claim 1; SEQ ID NO 3150; 56bp; English.

XX The invention relates to a protein comprising or having at least 50% identity to any of the 2469 amino acid sequences, identified in the CC specification (available on a computer readable format), or its fragment, CC expressed from 2469 of 2489 identified DNA coding regions from the CC Streptococcus pneumoniae type 4 strain genomic sequence appearing as CC AB556454. Also included are an antibody which binds one of the proteins, CC treating a patient by administering the protein, DNA or antibody (in a CC composition), a kit comprising first and second primers, which are the CC nucleic acid cited above or fragments between nucleotides 8-100 of a CC sequence not defined in the specification, for amplifying a target CC sequence contained within a Streptococcus nucleic acid sequence, where CC the first primer is substantially complementary to the target sequence CC and the second primer is substantially complementary to the complement of CC the target sequence, and where the parts of the primers having CC substantial complementarity define the termini of the target sequence to CC be amplified, assay comprising contacting a test compound with the CC protein, and determining whether the test compound binds to the protein CC and a Streptococcus pneumoniae bacterium, where one or more genes CC encoding the proteins has been rendered inactive. The proteins, nucleic CC acid molecules, antibody and compositions are useful as medicaments for CC treating or preventing a disease or infection due to streptococcus CC bacteria, particularly *S. pneumoniae*, such as pneumonia, sepsis, otitis CC media or ear infection. They are also useful in developing vaccines, CC diagnostics and antibiotics. The methods are useful for identifying CC immunodominant proteins. The present sequence is one of the 2469 proteins CC expressed by the identified coding regions from the genomic sequence. CC Note: The sequence data for this patent did not form part of the printed CC specification, but was obtained in electronic format directly from WIPO CC at fcp.wipo.int/pub/published_pct_sequences. (Updated on 23-Oct-2003 to CC standardise OS field)

XX Sequence 466 AA;

XX Query Match 42.4%; Score 407.5; DB 6; Length 466;

XX Best Local Similarity 45.0%; Pred. No. 4,8e-34;

XX Matches 94; Conservative 23; Mismatches 41; Indels 51; Gaps 7;

QY 3 DCDSTNSPLATATATLHNRGDIKIALFGNNAKTVANFVGLAQGTQDYSTONASGSPS 62
DB 277 DIRTV-EGPL-ATIKTNHGDLRIKLPFEHAPKTVANFVSL-SKD----- 317
QY 63 GPPYDGAVERVYQGMIOGDDPTGNGRGP---GYKPADBFHPILOFDKPYLLAMNAG 119
DB 318 -GYDGVTFRRILKDFMIQGGDPTGNGGSIYGSFEDBFSEBL-YNIRGALSMNAG 375
QY 120 PGTNGSQPFI-----TVGKTPHLNRRHTIFGEVIDAES 152
DB 376 PNTNGSQPFIYQNGHLPYSGKBITRGWPEPIAIBIYANGGCTPHLDNRHTVFGQLADEAS 435
QY 153 QRVVEAISKATDGNDRPTDPPVIESITI 181
DB 436 YAVLDAIAAVETGAMDKPVEDVIERIETI 464

XX RESULT 12

XX ABP81616

XX ID ABP81616 standard; protein; 466 AA.

XX AC ABP81616;

XX DT 04-MAR-2003 (first entry)

XX Streptococcus pneumoniae polypeptide SEQ ID NO 694.

XX Streptococcus pneumoniae; infection; otitis media; antibacterial;

XX diagnosis; gene therapy.
XX Streptococcus pneumoniae.

XX WO200283855-A2.

XX 24-OCT-2002.

XX 12-APR-2002; 2002WO-US011524.

XX 16-APR-2001; 2001US-0283948P.

XX 18-APR-2001; 2001US-0284433P.

XX (AMCV) AMERICAN CYANAMID CO.

XX Zagurebky RJ, Masi AW, Green BA, Chakravarti DN, Russell DP;

XX Wooters JL;

XX WPI; 2003-093010/08.

XX N-PSDB; ABZ42464.

XX New Streptococcus pneumoniae polynucleotides, useful for treating or

XX preventing *S. pneumoniae* infections, or non-systemic diseases, e.g.

XX otitis media, which are induced or exacerbated by *S. pneumoniae*.

XX Claim 42; Page 987-989; 1091pp; English.

XX The invention relates to isolated polynucleotides (ABZ72147-ABZ42522) of CC a Streptococcus pneumoniae genomic sequence, a fragment or degenerate CC variant of the polynucleotide or a nucleic acid sequence 95% identical to CC one of the polynucleotides. The *S. pneumoniae* polynucleotides and encoded CC polypeptides (ABP81299-ABP81674) are useful for treating or preventing *S. CC pneumoniae* infections or non-systemic diseases, e.g. otitis media, which CC are induced or exacerbated by *S. pneumoniae*. These are also useful for CC detecting *S. pneumoniae* in a biological sample or diagnosing *S. CC pneumoniae* infection in a subject. The polynucleotides have antibacterial CC activity and are useful in gene therapy

XX Sequence 466 AA;

XX Query Match 42.4%; Score 407.5; DB 6; Length 466;

XX Best Local Similarity 45.0%; Pred. No. 4,8e-34;

XX Matches 94; Conservative 23; Mismatches 41; Indels 51; Gaps 7;

QY 3 DCDSTNSPLATATATLHNRGDIKIALFGNNAKTVANFVGLAQGTQDYSTONASGSPS 62
DB 277 DIRTV-EGPL-ATIKTNHGDLRIKLPFEHAPKTVANFVSL-SKD----- 317
QY 63 GPPYDGAVERVYQGMIOGDDPTGNGRGP---GYKPADBFHPILOFDKPYLLAMNAG 119
DB 318 -GYDGVTFRRILKDFMIQGGDPTGNGGSIYGSFEDBFSEBL-YNIRGALSMNAG 375
QY 120 PGTNGSQPFI-----TVGKTPHLNRRHTIFGEVIDAES 152
DB 376 PNTNGSQPFIYQNGHLPYSGKBITRGWPEPIAIBIYANGGCTPHLDNRHTVFGQLADEAS 435
QY 153 QRVVEAISKATDGNDRPTDPPVIESITI 181
DB 436 YAVLDAIAAVETGAMDKPVEDVIERIETI 464

XX RESULT 13

XX ADM92196

XX ID ADM92196 standard; protein; 466 AA.

XX AC ADM92196;

XX DT 03-JUN-2004 (first entry)

XX *S. pneumoniae* antigenic protein sequence SegID393.

XX antibacterial; gene therapy; Streptococcus pneumoniae infection;

XX Streptococcus pneumoniae.

PN WO2004020609-A2.
XX 11-MAR-2004.
XX 02-SEP-2003; 2003WO-US027401.
XX 30-AUG-2002; 2002US-0407082P.
XX (TUFT) UNITV TUFTS.
XX
XX Cam1111 A, Hava DL;
XX WPI; 2004-239189/22.
XX N-PSDB; ADM91959.
XX
XX New Streptococcus pneumoniae nucleic acid molecules, useful for
PT diagnosing, treating and preventing active infections of Streptococcus
PT pneumoniae.
XX
XX Claim 8; SEQ ID NO 393; 123pp; English.
XX
XX This invention relates to novel isolated Streptococcus pneumoniae nucleic
CC acid molecules and the antigenic polypeptides encoded by them. The
CC invention may be useful for the production of compounds with an
CC antibacterial activity or for gene therapy. The nucleic acid molecules,
CC compositions and methods disclosed are useful for treating Streptococcus
CC pneumoniae infection. The present sequence is that of an S pneumoniae
CC protein of the invention.
XX
XX Sequence 466 AA;
SQ
Query Match 42.4%; Score 407.5; DB 8; Length 466;
Best Local Similarity 45.0%; Pred. No. 4.8e-34;
Matches 94; Conservative 23; Mismatches 41; Indels 51; Gaps 7;
QY 3 DCDSTVNSPLATATATLHTRGDIKIALFGNAPKTVANFVGLAOGTKDYSTONASGGS 62
DB 277 DIETV-EGPL-----ATIKTNHGDRLIKLFPFHAPKTVANFVSL-----SKD----- 317
QY 63 GPEYDGAVFHRYIQGFMIGSDPTGTGRGP---GKFADEFHPELOFDPKPYLLMANAG 119
DB 318 -GYDGVIFHRIIKDFMIGSDPTGTGMGESIYGSEFDEFSEEL-YNIRGALSWANAG 375
QY 120 PGTNSGQFPI-----TWGKTPHLNRHRTTIGEVIDAS 152
DB 376 PNTNSGQFPIVONQHLPYSKKEITRGWPEPIAEIYANOGTPTHLDRHRTVFGQLADEAS 435
QY 153 QRVEAISKATDGNDRPTDPVVISITTI 181
DB 436 YAVLDVIAAIVETGAMDKPVEDVVIETIEI 464
RESULT 14
ADK46830
ID ADK46830 standard; protein; 466 AA.
XX
XX ADK46830;
XX
XX 20-MAY-2004 (first entry)
XX
XX Streptococcus pneumoniae protein, Seq ID No 3345.
XX
XX Antibacterial; Gene therapy; Vaccine; Streptococcus pneumoniae.
XX
XX Streptococcus pneumoniae.
XX
XX US6699703-B1.
XX
XX 02-MAR-2004.
XX
XX 26-MAY-2000; 2000US-00583110.
XX
XX 02-JUL-1997; 97US-0051553P.
XX
XX

PR 12-MAY-1998; 98US-0085131P.
XX 30-JUN-1998; 98US-00107433.
XX
XX (GENO-) GENOME THERAPEUTICS CORP.
XX
XX Doucette-Stamm L, Bush D, Zeng Q, Opperman T, Housewart CE;
PI WPI; 2004-212399/20.
XX N-PSDB; ADK44169.
XX
XX New nucleic acid molecules and polypeptides useful for diagnosing,
PT preventing and treating pathological conditions resulting from bacterial
PT infection, e.g. Streptococcus pneumoniae infection, and in drug
PT screening.
XX
XX Disclosure; SEQ ID NO 3345; 301pp; English.
XX
XX The invention relates to isolated Streptococcus pneumoniae nucleic acids
CC and polypeptides. The nucleic acids and proteins are useful for
CC diagnosing, preventing and treating pathological conditions resulting
CC from bacterial infection, such as S. pneumoniae infection. These may also
CC be used for drug screening procedures. The present sequence represents a
CC Streptococcus pneumoniae polypeptide of the invention. Note: The sequence
CC data for this patent did not appear in the printed specification but was
CC obtained in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html.
XX
XX Sequence 466 AA;
SQ
Query Match 42.0%; Score 403.5; DB 8; Length 466;
Best Local Similarity 44.5%; Pred. No. 1.3e-33;
Matches 93; Conservative 23; Mismatches 42; Indels 51; Gaps 7;
QY 3 DCDSTVNSPLATATATLHTRGDIKIALFGNAPKTVANFVGLAOGTKDYSTONASGGS 62
DB 277 DIETV-EGPL-----ATIKTNHGDRLIKLFPFHAPKTVANFVSL-----SKD----- 317
QY 63 GPEYDGAVFHRYIQGFMIGSDPTGTGRGP---GKFADEFHPELOFDPKPYLLMANAG 119
DB 318 -GYDGVIFHRIIKDFMIGSDPTGTGMGESIYGSEFDEFSEEL-YNIRGALSWANAG 375
QY 120 PGTNSGQFPI-----TWGKTPHLNRHRTTIGEVIDAS 152
DB 376 PNTNSGQFPIVONQHLPYSKKEITRGWPEPIAEIYANOGTPTHLDRHRTVFGQLADEAS 435
QY 153 QRVEAISKATDGNDRPTDPVVISITTI 181
DB 436 YAVLDVIAAIVETGAMDKPVEDVVIETIEI 464
RESULT 15
ADR95835
ID ADR95835 standard; protein; 472 AA.
XX
XX ADR95835;
XX
XX 16-DEC-2004 (first entry)
XX
XX Novel S. pneumoniae protein sequence, SEQ ID 4470.
XX
XX Meningitis; bacteraemia; pneumonia; otitis media; vaccine;
XX bacterial infection.
XX
XX Streptococcus pneumoniae.
XX
XX US6800744-B1.
XX
XX 05-OCT-2004.
XX
XX 30-JUN-1998; 98US-00107433.
XX
XX 02-JUL-1997; 97US-0051553P.
XX
XX 12-MAY-1998; 98US-0085131P.
XX
XX

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: April 14, 2006, 17:19:08 / Search time 10.6721 Seconds
(without alignments)
1640.866 Million cell updates/sec

Title: US-10-620-246-12

Perfect score: 961
Sequence: 1 MADCDSVTNSPLATATATATLH.....ATDGNDRPTDPTVIESITTS 182

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	961	100.0	182	2	G70658	peptidylprolyl isoc
2	881	91.7	182	2	C86910	probable peptidyl
3	527.5	54.9	175	2	T51359	peptidylprolyl isoc
4	526.5	54.8	177	2	T36725	peptidylprolyl isoc
5	434	45.2	189	2	D84342	peptidylprolyl isoc
6	431.5	44.9	180	2	T43805	peptidylprolyl isoc
7	407.5	42.4	186	2	A95179	hypothetical prote
8	403.5	42.0	166	2	H98045	hypothetical prote
9	395	41.1	215	2	A71261	peptidylprolyl isoc
10	394.5	41.1	196	2	E86736	peptidylprolyl c1
11	380	39.5	629	2	T48940	hypothetical prote
12	374	38.9	160	2	C86151	peptidylprolyl isoc
13	358.5	37.3	155	2	T38930	peptidylprolyl isoc
14	351	36.5	610	2	T41399	probable cyclophil
15	350	36.4	169	2	T29283	peptidylprolyl isoc
16	339.5	35.3	197	2	C89662	hypothetical prote
17	336	35.0	194	2	AH1371	peptidylprolyl c1
18	336	35.0	194	2	AF1741	peptidylprolyl c1
19	334.5	34.8	201	2	T40819	peptidylprolyl isoc
20	333.5	34.7	164	2	H82708	peptidylprolyl isoc
21	332	34.5	164	2	C84777	peptidylprolyl isoc
22	332	34.5	193	2	E75543	peptidylprolyl isoc
23	331	34.4	471	1	S62590	peptidylprolyl c1
24	322.5	33.6	520	1	S64705	cyclophilin-like p
25	322	33.5	356	2	S62327	probable 40 kd pep
26	319.5	33.2	169	2	S22496	peptidylprolyl isoc
27	318	33.1	463	2	T39621	peptidylprolyl c1
28	315	32.8	212	2	A40047	peptidylprolyl isoc
29	314.5	32.7	350	2	E75341	peptidylprolyl c1

30	313.5	32.6	837	2	T27467	hypothetical prote
31	313	32.6	227	2	S71849	peptidylprolyl iso
32	310.5	32.3	145	1	CSYC42	peptidylprolyl iso
33	310.5	32.3	179	2	JT0686	peptidylprolyl iso
34	310.5	32.3	212	2	A54204	peptidylprolyl iso
35	309	32.2	222	2	T26500	peptidylprolyl iso
36	308	32.0	510	2	T05314	hypothetical prote
37	306	31.8	180	2	S51497	peptidylprolyl iso
38	305.5	31.6	176	2	F75261	peptidylprolyl iso
39	297.5	31.0	466	2	T18575	cyclophilin isoform
40	295.5	30.7	201	2	T18573	peptidylprolyl iso
41	295	30.7	147	2	T18577	peptidylprolyl iso
42	291.5	30.3	207	2	A40516	peptidylprolyl iso
43	291	30.3	208	1	CSH0B	peptidylprolyl iso
44	290.5	30.2	183	2	S71547	peptidylprolyl iso
45	290	30.2	223	1	CSNCM	peptidylprolyl iso

ALIGNMENTS

RESULT 1

G706598
peptidylprolyl isomerase (EC 5.2.1.8) p1r1 [similarity] - Mycobacterium tuberculosis (S
N)containing: cyclophilin
C/Species: Mycobacterium tuberculosis
C/Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C/Accession: G706598
R/Cole, S.T.; Broesch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon,
J.; Connor, R.; Davies, R.; Devlin, K.; Felwell, T.; Gentles, S.; Hamlin, N.; Holtroyd, S.
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A/Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A/Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A/Reference number: A70500; PMID:9825987; PMID:9634230
A/Accession: G706598
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-182 <COL>
A/Cross-references: UNIPROT:P71578; UNIPARC:UP1000013206D; GB:Z80233; GB:AL123456; NID:
A/Experimental source: strain H37Rv
C/Genetics:
A/Gene: p1r1
C/Suprafamily: peptidylprolyl isomerase; cyclophilin homology
C/Keywords: cis-trans-isomerase
F/9-182/Domain: cyclophilin homology <CYP>

Query Match 100.0%; Score 961; DB 2; Length 182;
Best local similarity 100.0%; Pred. No. 4,5e-80;
Matches 182; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MADCDSVTNSPLATATATATATLHNGDKIALFGNHAPKTVANFVGLQGTQDVSNTONASG 60
DB	1	MADCDSVTNSPLATATATATATLHNGDKIALFGNHAPKTVANFVGLQGTQDVSNTONASG 60
QY	61	PSGPFYDGAVERHVIQGFMIQGGDPTGTGSGGSGYFADEFFHDLQDKXYLLAMNAGP 120
DB	61	PSGPFYDGAVERHVIQGFMIQGGDPTGTGSGGSGYFADEFFHDLQDKXYLLAMNAGP 120
QY	121	GTNGSQPFTVVGKTPHNRHRTTGGVVIDAESQRYVEAISKTATDGNDRPTDPTVIESIT 180
DB	121	GTNGSQPFTVVGKTPHNRHRTTGGVVIDAESQRYVEAISKTATDGNDRPTDPTVIESIT 180
QY	181	IS 182
DB	181	IS 182

RESULT 2
C86910
probable peptidylprolyl cis-trans isomerase [imported] - Mycobacterium leprae
C/Species: Mycobacterium leprae
C/Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004

C/Accession: A71261
R/raeer, C.M.; Norris, S.J.; Weinstein, G.M.; White, O.; Sutton, G.G.; Dodson, R.; Gwin
rean, J.; Khaliq, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Uterback, T.; McD
they, L.; Weidman, J.; Smith, H.O.; Venter, J.C.
Science 281, 375-388, 1998
A/Title: Complete genome sequence of *Treponema pallidum*, the syphilis spirochete.
A/Reference number: A71250; M01D:98332770; PMID:9665876
A/Accession: A71261
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-215 <COL>
A/Cross-references: UNIPROT:Q66105; UNIPARC:UP10000132074; GB:AE001263; GB:AE000520; NIT
A/Experimental source: strain Nichols
C/Genetics:
A/Gene: TP0947
C/Superfamily: peptidylprolyl isomerase; cyclophilin homology
C/Keyword: cis-trans isomerase
F/34-206/Domain: cyclophilin homology <CYP>

Query Match 41.1%; Score 395; DB 2; Length 215;
Best Local Similarity 54.5%; Pred. No. 1.5e-28;
Matches 85; Conservative 15; Mismatches 44; Indels 12; Gaps 3;

QY 7 VNNSPLATAT--ATHTNRGDIKIALFGNHAFTYANFVGLAOGTKDYSTONASGSPSG 63
DB 29 VREGLAVADVGIYAVWETNRGTVLSLFFEKAPLTCNFFGLAEGT-----LAVCKGR 81
QY 64 PFPGDGVFHRVIOGFMIOGGDPGTGRGPGYKFADEFHPELOFDKPYLLMANAGPRTN 123
DB 82 PFQGLTFHRVLDKFMIOGGDPGNGTGPYQPPDCDPLALNHSDBGLSMANAGPRTN 141
QY 124 GSQFFITVGTCTPHANRRHTIFGEVIDAESQRYVEAI 159
DB 142 GSQFFITVATPWLDDGHTVFGKV--EGMEVVAHAI 175

RESULT 10

E86736
peptidyl-prolyl cis-trans isomerase [imported] - *Lactococcus lactis* subsp. *lactis* (strain
C/Species: *Lactococcus lactis* subsp. *lactis*
C/Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 09-Jul-2004
C/Accession: E86736
R/Bolotin, A.; Wincker, P.; Manger, S.; Jallion, O.; Malarme, K.; Weissenbach, J.; Ehrlich
Genome Res. 11, 731-753, 2001
A/Title: The complete genome sequence of the lactic acid bacterium *Lactococcus lactis* s
A/Reference number: A86625; M01D:21235186; PMID:11337471
A/Accession: E86736
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-196 <STO>
A/Cross-references: UNIPROT:Q9CH46; UNIPARC:UP1000006930; GB:AE005176; PID:g12723824; F
A/Experimental source: strain IL403
C/Genetics:
A/Gene: PPIB
C/Superfamily: peptidylprolyl isomerase; cyclophilin homology

Query Match 41.1%; Score 394.5; DB 2; Length 196;
Best Local Similarity 43.1%; Pred. No. 1.4e-28;
Matches 84; Conservative 26; Mismatches 40; Indels 45; Gaps 4;

QY 16 TATLTNRGDIKIALFGNHAFTYANFVGLAOGTKDYSTONASGSPSGPYDGAVFHRVY 75
DB 15 TATLTNRGDIKIALFGNHAFTYANFVGLAOGTKDYSTONASGSPSGPYDGAVFHRVY 59
QY 76 QGFMTIOGSPGTGRGSP--GYKPADERHPELOFDKPYLLMANAGPRTNNGSPFT--- 129
DB 60 KDFMTIOGSPGTGRGSP--GYKPADERHPELOFDKPYLLMANAGPRTNNGSPFT--- 118
QY 130 -----TVGKTPIHNRHTIFGEVIDAESQRYVEAI 166
DB 119 KULPYKSLIOGSPGEEVALETYGGTPIHNRHTIFGEVIDAESQRYVEAI 178
QY 167 NDRPTDPPVIESITTI 181

DB 179 QDKPVDVIVIESVEI 193

RESULT 11

T48940
hypothetical protein F14L2.150 - *Arabidopsis thaliana*
C/Species: *Arabidopsis thaliana* (mouse-ear cress)
C/Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 09-Jul-2004
C/Accession: T48940
R/Jordan, N.; Bangert, S.; Wiedemann, R.; Voese, H.; Uneseld, M.; Mewes, H.W.; Rudd, S.;
submitted to the Protein Sequence Database, April 2000
A/Reference number: Z25008
A/Accession: T48940
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-629 <TOR>
A/Cross-references: UNIPROT:Q9LXN7; UNIPARC:UP1000004046D; EMBL:AL353818; GSPDB:GN00061;
A/Experimental source: cultivar Columbia, BAC clone F14L2
C/Genetics:
A/Gene: ATSP.F14L2.150
A/Map position: 3
A/Intons: 62/2; 91/1; 120/3; 200/3; 241/2; 266/3; 302/3; 361/3; 434/2; 476/3; 551/2; 60

Query Match 39.5%; Score 380; DB 2; Length 629;
Best Local Similarity 47.6%; Pred. No. 1.3e-26;
Matches 79; Conservative 21; Mismatches 46; Indels 20; Gaps 4;

QY 19 LHTNRGDIKIALFGNHAFTYANFVGLAOGTKDYSTONASGSPSGPYDGAVFHRVIOGP 78
DB 478 MHTTLDGDIHMKLYPBECPPTVENF-----TTHCRNG---YDNLHFRVIRGR 522
QY 79 MIOGDPGTGRGSP--GYKPADERHPELOFDKPYLLMANAGPRTNNGSPFTVGTTP 135
DB 523 MIOGDPDLDCGTGCGSIMGKREDFERHKSRLRDRPTLSMANAGPRTNNGSPFTVGTTP 582
QY 136 HLNRHTIFGEVIDAESQRYVEAISTATGNDPPTVIESITTI 181
DB 583 WLDNKTIVFGRVV--KMDVVOGIEKVTDKDRPYQVKILNVTV 626

RESULT 12

C86151
P22M8.7 protein - *Arabidopsis thaliana*
C/Species: *Arabidopsis thaliana* (mouse-ear cress)
C/Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C/Accession: C86151
R/Theologis, A.; Ecker, J.R.; Palm, C.T.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Coma, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
anzen, N.F.; Hughes, B.; Hultzer, L.
Nature 408, 816-820, 2000
A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kilm,
C.A.; Li, J.H.; Li, J.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lurie, J.S.; Malt, R.; Marshall,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A/Title: Sequence and analysis of chromosome 1 of the plant *Arabidopsis*.
A/Reference number: A86141; M01D:21016719; PMID:11130712
A/Accession: C86151
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-160 <STO>
A/Cross-references: UNIPROT:Q9LPC7; UNIPARC:UP10000008B15; GB:AE005172; NID:g8570445; PI

C/Superfamily: peptidylprolyl isomerase; cyclophilin homology

Query Match 38.9%; Score 374; DB 2; Length 160;
Best Local Similarity 47.3%; Pred. No. 8.2e-27;
Matches 80; Conservative 22; Mismatches 47; Indels 20; Gaps 4;
QY 16 TATLTNRGDIKIALFGNHAFTYANFVGLAOGTKDYSTONASGSPSGPYDGAVFHRVY 75

D_b 2 SVTLATNTAGDICKCEIFCDEVPKSAENP₁ALC-----ASG-----YYDGI₂TFH₃NI 46

O_Y 76 OGFMIOGGDPPTGTGSGP---GYKPAD₁EFH₂PELO₃FDKDYLLAMANA₄SGTNGS₅OFF₆ITV₇G 132

D_b 47 KGFMIQGGDPKGTGTGCGTGISIMCKEN₁DEIR₂BSL₃KHNAG₄GLMS₅ANGSP₆INT₇NSO₈PFIT₉YA 106

O_Y 133 KTFPHNRHNTTFGEYIDNASQRV₁EAISK₂TATD₃GN₄DR₅PTDP₆VIES₇IT₈I 181

D_b 107 KQPHLNGLYTTIFGKVI--HGFEVL₁DI₂ME₃TQ₄GG₅PD₆RL₇PA₈EL₉RL₁₀RV₁₁I 153

```

RESULT 13
T38930
peptidylprolyl isomerase (EC 5.2.1.8) SPAC57A10.03 [similarity] - fission yeast (Schizosaccharomyces pombe)
N:Contains: cyclophilin
C:Species: Schizosaccharomyces pombe
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C:Accession: T38930
R:Badcock, K.; Churher, C.M.; Wood, V.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, April 1997
A:Reference number: Z21818
A:Accession: T38930
A:Status: preliminary; translated from GB/EMBL/DBD
A:Molecule type: DNA
A:Residues: 1-155 <BAD>
A:Cross-references: UNIPROT:P87051; UNIPARC:UP10000128C69; EMBL:Z94864; PIDD:CAB08166.1;
A:Experimental source: strain 972h-; cosmid c57A10
C:Genetics:
A:Gene: SPDB:SPAC57A10.03
A:Map position: 1
A:Intons: 10/3; 24/3; 42/3; 69/2; 101/3
C:Superfamily: peptidylprolyl isomerase; cyclophilin homology
C:Keywords: cis-trans-isomerase
;1-155/Domain: cyclophilin homology <CYP>

```

Query Match	37.3%	Score 358.5	DB 2	Length 155
Best Local Similarity	45.6%	Pred. No. 26-25		
Matches	78	Conservative	23	Mismatches 49
			Indels	21
			Gaps	4

QY	15	ATATLHTNRGDIKIALPFGNHAPKTVANFVGLAQGTQDYS	TONASGSPGSPFYDGA	VFRV	74
Db	2	ANVELQTSGLKILIELYTEHAPKTCQNFYTLAK	-----	EGYDVGI	FRV
					46
QY	75	IQGFMIQGGDPYTCGRGSP	---GYKPADEPHELPD	KPYLLANNAAGP	NGSFFITV
					131
Db	47	IPDVPYIQGGDPYTCGRGGS	IYGDKPDDEIHSDLHT	AGAGLISMANA	PNTNNSSQ
					106
QY	132	GKTEPLNRRHRTFGEVIDASQ	RVEAISKATDGNDR	PTDPV	VIYSITI
					181
Db	107	APYMLDCKHITFERVVSGLS	-VCKRNGLIRTSDD	PIEELKTIK	IKVAL
					155

RESULT 14
 T41399
 Probable cyclophilin-related peptidyl prolyl cis-trans isomerase - fission yeast (Schizosaccharomyces pombe)
 C:Species: Schizosaccharomyces pombe
 C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
 C:Accession: T41399
 R:Wood, V.; Rajandream, M.A.; Barrell, B.G.; Murphy, L.; Harris, D.
 Submitted to the EMBL Data Library, May 1998
 A:Reference number: Z21991
 A:Accession: T41399
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-610 <MO>
 A:Cross-references: UNIPROT:O74942; UNIPARC:UPI0000069EF1; EMBL:AL023704; PIDD:CAA19257
 A:Experimental source: strain 972h; cosmid c553
 C:Genetics:
 A:Gene: SPDB:SPDC553.04
 A:Map position: 3
 A:Introns: 24/1
 A:449-608/Domain: cyclophilin homology <CP>

Query March	36.5%;	Score 351;	DB 2;	Length 610;
Best Local Similarity	42.3%;	Pred. No. 5,3e-24;		
Matches	80;	Conservative	22;	Mismatches 53; Indels 34; Gaps 6
Qy	7	VTNSP-----LATAATLHTNRGDIKTALFQNHAKPTVYANFVGLAOGTKDYSTONAS	58	
Db	439	VTNNIEGRQENGNMILGKRAAIHTHTQGDISIDLYEEBAKAVONF-----TTAAE	487	
Qy	59	GGSGPFPYDAVFHRVYOGFMIOGGDDPTGTRGSPGY---FEADBFHPELOFDKPYLLAM	115	
Db	488	NG----YYDNTIFRRIIKNFMIQGGDELDGDTGSESTWKQDFEBSISNLKHDRPFTYSM	543	
Qy	116	ANAGPFGNSGOFITVVGKTPHLNRRHTTFGEV---IDAQSRYVEAISKTATDGNDRPTD	172	
Db	544	ANSQPNNTNGSOFITTDLTLPWLDGKHTIFAAVAYGLD-----VHRIEGSTDKYDRPLE	598	
Qy	173	PVVEISITIT	181	
Db	599	PTKINISTIT	607	

RESULT 15
T29283
peptidylprolyl isomerase (EC 5.2.1.8) C34D4.12 [similarity] - *Caenorhabditis elegans*
N/Contains: cyclophilin
C/Species: *Caenorhabditis elegans*
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C/Accession: T29283
R/Du. Z.; Le, T.T.
submitted to the EMBL Data Library, May 1996
A/Description: The sequence of C. elegans cosmid C34D4.
A/Reference number: Z20600
A/Accession: T29283
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residue 1-169 <DDZ>
A/Cross-references: UNIPROT:Q18445; UNIPARC:UPI000007D84B; EMBL:U58755; PIDD:AA80697.1
A/Experimental source: strain Bristol N2; clone C34D4
C/Genetics:
A/Gene: CESP:C34D4.12
A/Map position: 4
A/Intons: 69/1; 102/1
C/Superfamily: peptidylprolyl isomerase; cyclophilin homology
C/Keywords: cis-trans-isomerase

Query Match	36.4%	Score 350;	DB 2;	Length 169;
Best Local Similarity	47.1%;	Pred. No. 1.3e-24;		
Matches	73;	Conservative	20;	Mismatches 42;
			Indels	20;
			Gaps	3.
QY	19	LHTRRGDKIALPEGNHAKPTVIANFVGIAOGTKDYSTONASGSGSPGFYDGAVFHRVIOGF	78	
Db	14	LDITTMGMIABEIVYNNHAPRTCCQNSQLAKN-----YNGNGIIFRIINDF	58	
QY	79	MIOGGDPTGTGRCGP---GYKFADEFHEPELOFDKPYLLAMANNAGDPTNGSOFPTTVGKTP	135	
Db	59	MIOGGDPTGTGRCGASLYGDKPSDEIDEBRLKHTGAGILISMANNAGDPTNGSOFPTTLAPTO	118	
QY	136	HLNRHRTIPEGEVIDABESQRVEALSKATGNDNDP	170	
Db	119	HLDSKRTIFGRV--AAGMKVIANNGRVDTNNHDP	151	

Search completed: April 14, 2006, 17:34:32
Job time : 12.6721 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 14, 2006, 17:04:10 ; Search time 67.152 Seconds

(without alignments)
1912.171 Million cell updates/sec

Title: US-10-620-246-12

Perfect score: 961

Sequence: 1 MADCDSVTMSPLATATATLH.....ATGNDPRDPVIESITTS 182

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : UniProt_05.80.*

1: uniProt_sprot.*

2: uniProt_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	961	100.0	182	1 PPIA_MYCBO	P65763 mycobacteri
2	961	100.0	182	1 PPIA_MYCTU	P65762 mycobacteri
3	881	91.7	182	1 PPIA_MYCLB	Q96d69 mycobacteri
4	877	91.3	182	2 Q744Q3_MYCPA	Q4jy65 corynebacte
5	717.5	74.7	174	2 Q4JYC5_CORJX	Q4jy65 corynebacte
6	716	74.5	174	2 Q6M8Y0_CORGL	Q6m8y5 corynebacte
7	716	74.5	187	2 Q6NDA5_CORGL	Q6n8y5 corynebacte
8	710	73.9	180	2 Q6NKJ3_CORDI	Q6nkj3 corynebacte
9	708	73.7	202	2 Q6FUI9_CORBP	Q6fui9 corynebacte
10	611	63.6	170	2 Q6A9U1_PROAC	Q6a9u1 propionibac
11	543.5	56.6	186	2 Q82FC6_STRAW	Q82fc6 streptomyce
12	528	54.9	179	2 Q6AHM5_LEIKX	Q6ahm5 leishonia x
13	527.5	54.9	175	1 PPIB_STRCH	P77949 streptomyce
14	526.5	54.8	177	2 Q9XA08_STRCO	Q9xa08 streptomyce
15	503.5	52.4	181	2 Q4NK35_9MICC	Q4nk35 streptomyce
16	465.5	48.4	182	2 Q8GAB6_BIFIO	Q8g466 bifidobacte
17	464	48.3	209	2 Q5UKX8_HAAMA	Q5ukx8 halosarcila
18	460.5	47.9	188	2 Q5X3T8_LEGPA	Q5x3t8 legionella
19	454.5	47.3	188	2 Q5XW81_LEGPA	Q5xw81 legionella
20	454.5	47.3	188	2 Q5ZU46_LEGPH	Q5zu46 legionella
21	447.5	46.6	211	2 Q6MRB4_BDEBA	Q6mr46 bdellovibri
22	434	45.2	189	2 Q9HNW6_HALSA	Q9hnw6 halobacteri
23	431.5	44.9	180	2 Q50586_HALSA	Q50586 halobacteri
24	429.5	44.7	188	2 Q4NTJ5_SDELT	Q4ntc5 anaeromyxob
25	408	42.5	186	2 Q7X1L9_9BACT	Q7x1l9 leptospirill
26	407.5	42.4	178	2 Q97PR4_STRPN	Q97pr4 streptococc
27	406.5	42.3	322	2 Q73JMG_TRBDP	Q73jmg streptococc
28	406	42.2	466	2 Q8K6F2_STRP3	Q8k6f2 streptococc
29	406	42.2	470	2 Q879E7_STRP3	Q879e7 streptococc
30	403.5	42.0	466	2 Q8CTU1_STRR6	Q8cty1 streptococc
31	401.5	41.8	143	2 Q7UP01_RHOBA	Q7up01 rhodospirell

32	398.5	41.5	253	2 Q5BXN7_SCHTA	Q5bxn7 schistosoma
33	397	41.3	470	2 Q5XAQ1_STRP6	Q5xaq1 streptococc
34	395	41.1	215	1 PPIB_TREPA	Q66i05 treponema p
35	395	41.1	344	2 Q6AMQ1_DESPS	Q6amq1 desulfocale
36	394.5	41.1	196	2 Q9CH46_LACLA	Q9ch46 lactococcus
37	394	41.0	471	2 Q8DVJ7_STRMU	Q8dvj7 streptococc
38	393	40.9	466	2 Q8B1N1_STRAS	Q8b1n1 streptococc
39	392	40.8	470	2 Q8P009_STRF8	Q8p009 streptococc
40	388.5	40.4	468	2 Q5M0N3_STRT1	Q5m0n3 streptococc
41	388.5	40.4	468	2 Q5M569_STRT2	Q5m569 streptococc
42	388	40.4	466	2 Q8E747_STRJ3	Q8e747 streptococc
43	387	40.3	159	2 Q6MC42_PARIW	Q6mc42 parachlamyd
44	386	40.2	181	2 Q7XXQ2_ORYSA	Q7xxq2 oryza sativ
45	386	40.2	423	2 Q6ZJ09_ORYSA	Q6zj09 oryza sativ

ALIGNMENTS

RESULT 1
PPIA_MYCBO STANDARD; PRT; 182 AA.
ID PPIA_MYCBO
AC P65763; P1578;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Probable peptidyl-prolyl cis-trans isomerase A (EC 5.2.1.8) (PPIase A)
DE (Rotamase A).
GN Name=PPIA; Ordered locus names=Mb0009;
OS Mycobacterium bovis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacteriaceae; Mycobacteriaceae; Mycobacterium;
OC Mycobacterium tuberculosis complex.
OX NCBI_TaxID=1765;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=AF2122/97;
RX MEDLINE=22709107; PubMed=12788972; DOI=10.1073/pnas.1130426100;
RA Garnier T., Eglmeier K., Camus J.-C., Medina N., Mansoor H.,
RA Pryor M., Dutfoy S., Grondin S., Lacroix C., Monsemp C., Simon S.,
RA Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,
RA Parkhill J., Barrell B.G., Cole S.T., Gordon S.V., Hewinson R.G.,
RT "The complete genome sequence of Mycobacterium bovis."
RL Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882(2003).
CC -1- FUNCTION: PPIases accelerate the folding of proteins. It catalyzes
the cis-trans isomerization of proline imidic peptide bonds in
oligopeptides (By similarity).
CC -1- CATALYTIC ACTIVITY: Peptidylproline (omega=180) = peptidylproline
(omega=0).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Belongs to the cyclophilin-type PPIase family.
CC -1- SIMILARITY: Contains 1 PPIase cyclophilin-type domain.
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use as long as its content is in no way modified and this statement is not
removed.
CC EMBL; BX248334; CAD92871.1; -; Genomic DNA.
CC HSSP; P52013; IHOP.
CC SNR; P65763; 12-182.
CC InterPro; IPR002130; CSA_PPIase.
CC Pfam; PF00160; Pro_isomerase; 1.
CC PRINTS; PR00153; CSAPISWASR.
CC PROSITE; PS00170; CSA_PPIASE_1; 1.
CC PROSITE; PS50072; CSA_PPIASE_2; 1.
CC Complete proteome; Isomerase; Rotamase.
CC DOMAIN 13 181 PPIase cyclophilin-type.
CC SQ SEQUENCE 182 AA; 19239 MW; B3D83E7F946B8CA CRC64;
Query Match 100.0%; Score 961; DB 1; Length 182;
Best Local Similarity 100.0%; Pred. No. 3,7e-84;

```

Matches 182; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MADCSVTNSPLATATATHTNRGDIKIALFGNHAKPTVAVVGLAQGTQDYSTONASGG 60
DB 1 MADCSVTNSPLATATATHTNRGDIKIALFGNHAKPTVAVVGLAQGTQDYSTONASGG 60
QY 61 PGGPFYDGVAFHVRVIOGFMIQGGDPGTGRGGPGYKFADEPHDELQFDKPYLLAMANNAGP 120
DB 61 PGGPFYDGVAFHVRVIOGFMIQGGDPGTGRGGPGYKFADEPHDELQFDKPYLLAMANNAGP 120
QY 121 GTNGSQFFITVGTGTPHLNRHRTIFGEVIDAESQVVEAISKATDGNDRPTDPTVIESIT 180
DB 121 GTNGSQFFITVGTGTPHLNRHRTIFGEVIDAESQVVEAISKATDGNDRPTDPTVIESIT 180
QY 181 IS 182
DB 181 IS 182

RESULT 2
PPIA MYCTU STANDARD; PRT; 182 AA.
ID P65762; P71578;
AC 01-NOV-1997 (Rel. 35, Created)
DT 25-OCT-2004 (Rel. 45, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Probable peptidyl-prolyl cis-trans isomerase A (EC 5.2.1.8) (PPIase A)
(Genbank A).
GN Name=ppia; OrderedLocusNames=Rv0009, MT0011; ORFNames=MT0011.08;
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacteriaceae; Mycobacteriaceae; Mycobacterium;
OC Mycobacterium tuberculosis complex.
NCBI_TaxId=1773;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=H37RV;
RX MEDLINE=98295987; PubMed=9634230; DOI=10.1038/31159;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C.M.,
RA Harris D.B., Gordon S.V., Eigemeier K., Gas S., Barry C.E. III,
RA Tekala F., Badcock K., Basham D., Brown D., Chillingworth T.,
RA Connor R., Davies R.M., Devlin K., Felwell T., Gentles S., Hamlin N.,
RA Holroyd S., Hornsby T., Jagels K., Krogan A., McLean J., Moule S.,
RA Murphy L.D., Oliver S., Osborne J., Quail M.A., Rajandream M.A.,
RA Rogers J., Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Suleron J.E., Taylor K., Whitehead S., Barrett B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence."
RT Nature 393:537-544 (1998).
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=CDC 1551 / Oshkosh;
RX MEDLINE=22206484; PubMed=12218036;
DOI=10.1128/JB.184.19.5479-5490.2002;
RA Felschmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J.D., Deboy R.T., Dodson R.J., Gwin M.L., Hat D.H.,
RA Hickney B.K., Kolony J.F., Nelson W.C., Umayan L.A., Ermolaeva M.D.,
RA Salzberg S.L., Delcher A., Uitterback T.R., Weidman J.F., Kouri H.M.,
RA Gill J., Mikula A., Bishai W., Jacobs W.R. Jr., Venter J.C.,
RA Fraser C.M.;
RT "Whole-genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains."
RT J. Bacteriol. 184:5479-5490 (2002).
CC -1- FUNCTION: PPIases accelerate the folding of proteins. It catalyzes
CC the cis-trans isomerization of proline imide peptide bonds in
CC oligopeptides (By similarity).
CC -1- CATALYTIC ACTIVITY: Peptidylproline (omega=180) = peptidylproline
CC (omega=0).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Belongs to the cyclophilin-type PPIase family.
CC -1- SIMILARITY: Contains 1 PPIase cyclophilin-type domain.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration

```

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
DR EMBL: BX842572; CAB02430.1; -; Genomic DNA.
DR EMBL: AB000516; AK44233.1; -; Genomic DNA.
DR PIR: G70698; G70698.
DR PDB: 1W74; X-ray; A/B=2-182.
DR TIGR: MT0011; -.
DR Tuberculise; Rv0009; -.
DR InterPro: IPR002130; CSA_PPIase.
DR Pfam: PF00160; Pro_Isoomerase; 1.
DR PRINTS: PR00153; CSAPPISRRASE.
DR PROSITE: PS00170; CSA_PPIASE_1; 1.
DR PROSITE: PS50072; CSA_PPIASE_2; 1.
KW 3D-structure; Complete proteome; Isomerase; Rotamase.
FT DOMAIN 13 181 PPIase cyclophilin-type.
SQ SEQUENCE 182 AA; 19239 MW; B3DB3E7F9486BFCA CRC64;

Query Match 100.0%; Score 961; DB 1; Length 182;
Best Local Similarity 100.0%; Pred. No. 3.7e-84;
Matches 182; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MADCSVTNSPLATATATHTNRGDIKIALFGNHAKPTVAVVGLAQGTQDYSTONASGG 60
DB 1 MADCSVTNSPLATATATHTNRGDIKIALFGNHAKPTVAVVGLAQGTQDYSTONASGG 60
QY 61 PGGPFYDGVAFHVRVIOGFMIQGGDPGTGRGGPGYKFADEPHDELQFDKPYLLAMANNAGP 120
DB 61 PGGPFYDGVAFHVRVIOGFMIQGGDPGTGRGGPGYKFADEPHDELQFDKPYLLAMANNAGP 120
QY 121 GTNGSQFFITVGTGTPHLNRHRTIFGEVIDAESQVVEAISKATDGNDRPTDPTVIESIT 180
DB 121 GTNGSQFFITVGTGTPHLNRHRTIFGEVIDAESQVVEAISKATDGNDRPTDPTVIESIT 180
QY 181 IS 182
DB 181 IS 182

RESULT 3
PPIA MYCLE STANDARD; PRT; 182 AA.
ID Q9CDE9;
AC 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Probable peptidyl-prolyl cis-trans isomerase A (EC 5.2.1.8) (PPIase A)
(Rotamase A).
GN Name=ppia; OrderedLocusNames=ML0011;
OS Mycobacterium leprae.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.
NCBI_TaxId=1765;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=TN;
RX MEDLINE=21128732; PubMed=11234002; DOI=10.1038/35059006;
RA Cole S.T., Eigemeier K., Parkhill J., James K.D., Thomson N.R.,
RA Wheeler P.R., Honore N., Garnier T., Churcher C.M., Harris D.E.,
RA Mungall K.L., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R.M., Devlin K., Duthoy S., Felwell T., Fraser A., Hamlin N.,
RA Holroyd S., Hornsby T., Jagels K., Lacroix C., Maclean J., Moule S.,
RA Murphy L.D., Oliver K., Quail M.A., Rajandream M.A., Rutherford K.M.,
RA Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R.,
RA Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,
RA Barrett B.G.;
RT "Massive gene decay in the leprosy bacillus."
RT Nature 409:1007-1011 (2001).
CC -1- FUNCTION: PPIases accelerate the folding of proteins. It catalyzes
CC the cis-trans isomerization of proline imide peptide bonds in
CC oligopeptides (By similarity).
CC -----

```

```

CC -1- CATALYTIC ACTIVITY: Peptidylproline (omega=180) = peptidylproline
CC (omega=0).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Belongs to the cyclophilin-type Ppiase family.
CC -1- SIMILARITY: Contains 1 ppiase cyclophilin-type domain.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
DR EMBL, AL583917, GAC29519.1, -, Genomic_DNA.
DR PIR, C86910, C86910.
DR HSSP, P52013, 1HOP.
DR SMR, Q9CDE9, 12-182.
DR Lepidoma, ML0011, -.
DR InterPro, IPR002130, CSA_Ppiase.
DR Pfam, PF00160, Pro_Isomerase, 1.
DR PRINTS, PR00153, CSAPPISMASE.
DR PROSITE, PS00170, CSA_PPIASE_1, 1.
DR PROSITE, PS50072, CSA_PPIASE_2, 1.
KW Complete proteome; Isomerase; Rotamase.
PT DOMAIN 13
SQ SEQUENCE 182 AA; 19367 MW; 6A6D804348A9544 CRC64;

Query Match 91.7%; Score 881; DB 1; Length 182;
Best Local Similarity 90.1%; Pred. No. 1,8e-76;
Matches 164; Conservative 9; Mismatches 9; Indels 0; Gaps 0;

QY 1 MADCDSTNSPLATATLTHTNRGDIKIALFGNHAFTVAVFVGLAQGTQDYSTONASGG 60
DB 1 MADCDSTNSPLATATLTHTNRGDIKIALFGNHAFTVAVFVGLAQGTQDYSTONASGG 60
QY 61 PSGPFDGAVFHHVIRQFMIOGSDPTGTGSGPGYKPADSFHPELOFDKXYLLAMANAAGP 120
DB 61 PSGPFDGAVFHHVIRQFMIOGSDPTGTGSGPGYKPADSFHPELOFDKXYLLAMANAAGP 120
QY 121 GTNGSQFFITVGKTPHLNRHTIFGEVIDAESQRVVEAISKTATDGNDRPTDPVVISIT 180
DB 121 GTNGSQFFITVGKTPHLNRHTIFGEVIDAESQRVVEAISKTATDGNDRPTDPVVISIT 180
QY 181 IS 182
DB 181 IS 182

RESULT 4
Q74403 MYCPA PRELIMINARY; PRT; 182 AA.
ID Q74403 MYCPA PRELIMINARY; PRT; 182 AA.
AC Q74403;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Pp1A.
GN Name=Pp1A; OrderedLocustNames=MAP0011;
OS Mycobacterium paratuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacteriaceae; Mycobacteriaceae; Mycobacterium;
OC Mycobacterium avium complex (MAC).
OX NCBI_TaxID=1770;
OX [1] _
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=K10;
RA Li U., Bannantline J., Zhang Q., Amonsik A., Alt D., Kapur V.;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Contains 1 Ppiase cyclophilin-type domain.
DR EMBL, AB017227, AA002328.1, -, Genomic_DNA.
DR HSSP, Q27450, 1A58.
DR SMR, Q74403, 14-182.
DR GO, GO:0016853, P:isomerase activity; IEA.
DR GO, GO:0003755, P:peptidyl-prolyl cis-trans isomerase activity; IEA.
DR GO, GO:0006457, P:protein folding; IEA.

```

```

DR InterPro, IPR002130, CSA_Ppiase.
DR Pfam, PF00160, Pro_Isomerase, 1.
DR PRINTS, PR00153, CSAPPISMASE.
DR PROSITE, PS00170, CSA_PPIASE_1, 1.
DR PROSITE, PS50072, CSA_PPIASE_2, 1.
KW Complete proteome; Isomerase; Rotamase.
SQ SEQUENCE 182 AA; 19327 MW; 7FBB40307091A12E CRC64;

Query Match 91.3%; Score 877; DB 2; Length 182;
Best Local Similarity 89.6%; Pred. No. 4,4e-76;
Matches 163; Conservative 11; Mismatches 8; Indels 0; Gaps 0;

QY 1 MADCDSTNSPLATATLTHTNRGDIKIALFGNHAFTVAVFVGLAQGTQDYSTONASGG 60
DB 1 MADCDSTNSPLATATLTHTNRGDIKIALFGNHAFTVAVFVGLAQGTQDYSTONASGG 60
QY 61 PSGPFDGAVFHHVIRQFMIOGSDPTGTGSGPGYKPADSFHPELOFDKXYLLAMANAAGP 120
DB 61 PSGPFDGAVFHHVIRQFMIOGSDPTGTGSGPGYKPADSFHPELOFDKXYLLAMANAAGP 120
QY 121 GTNGSQFFITVGKTPHLNRHTIFGEVIDAESQRVVEAISKTATDGNDRPTDPVVISIT 180
DB 121 GTNGSQFFITVGKTPHLNRHTIFGEVIDAESQRVVEAISKTATDGNDRPTDPVVISIT 180
QY 181 IS 182
DB 181 IS 182

RESULT 5
Q4JUC5 CORJK PRELIMINARY; PRT; 174 AA.
ID Q4JUC5 CORJK PRELIMINARY; PRT; 174 AA.
AC Q4JUC5;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Peptidyl-prolyl cis-trans isomerase (EC 5.2.1.8).
GN Name=Pp1A; ORFNames=jK0033;
OS Corynebacterium jeikeium (strain K411).
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacteriaceae; Corynebacteriaceae; Corynebacterium.
OX NCBI_TaxID=306537;
OX [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=K411;
RX PubMed=15968079; DOI=10.1128/JB.187.13.4671-4682.2005;
RA Tsuchi A., Kaiser O., Hain T., Goessmann A., Weisshaar B.,
RA Albertmeier A., Bekel T., Bischoff N., Brune I., Chakraborty T.,
RA Kallinowski J., Meyer F., Rupp O., Schneider S., Viehoveer P.,
RA Puhler A.;
RT "Complete Genome Sequence and Analysis of the Multiresistant
RT Nosocomial Pathogen Corynebacterium jeikeium K411, a Lipid-Requiring
RT Bacterium of the Human Skin Flora.";
RL J. Bacteriol. 187:4671-4682(2005).
RL [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=K411;
RA Ilnke B., Tsuchi A.;
RL Submitted (DEC-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL, CR31997, CA136182.1, -, Genomic_DNA.
KW Isomerase.
SQ SEQUENCE 174 AA; 18825 MW; B72873CD42C381B6 CRC64;

Query Match 74.7%; Score 717.5; DB 2; Length 174;
Best Local Similarity 77.7%; Pred. No. 9,2e-61;
Matches 136; Conservative 12; Mismatches 24; Indels 3; Gaps 1;

QY 7 VTNSPLATATLTHTNRGDIKIALFGNHAFTVAVFVGLAQGTQDYSTONASGGSPGPFY 66
DB 1 MTNK---TATLHTNGSDIALDLFGNHAFTVAVFVGLAQGTQDYSTONASGGSPGPFY 57
QY 67 DGAVFHHVIRQFMIOGSDPTGTGSGPGYKPADSFHPELOFDKXYLLAMANAAGPQTNGSQ 126

```

```

Db      58 DGAIFRIIDGFMIOGGDPPTGTRGCGPGYGPDEHFPELOFDPFPLAAMNAGPGTNGSQ 117
Qy      127 PFITVGTPLNRRHTIFGEVIDAESQRVVEAISKTATDGNDRPTDPTVIESITTI 181
      118 PFITVATPLNRRHTIFGEVTDKESQKVAKISRVATDMDRPNDDVIESVEI 172

RESULT 6
Q6M8Y0 CORGL PRELIMINARY; PRT; 174 AA.
ID Q6M8Y0;
AC Q6M8Y0;
DT 05-JUL-2004 (T-EMBLrel. 27, Created)
DT 05-JUL-2004 (T-EMBLrel. 27, Last sequence update)
DE 05-JUL-2004 (T-EMBLrel. 27, Last annotation update)
DE PEPTIDYL-PROLYL CIS-TRANS ISOMERASE B (EC 5.2.1.8).
GN Name:pplA; OrderedlocusNames=cg0048;
OS Corynebacterium glutamicum (Brevibacterium flavum).
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacteriaceae; Corynebacteriaceae; Corynebacterium.
OX NCBI_TaxID=1718;
(1) _SEQUENCE.
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;
RC PubMed=12948625; DOI=10.1016/S0168-1656(03)00154-8;
RA Kallinowski J., Bathe B., Bartels D., Bischoff N., Boett M.,
RA Burkoverki A., Dusch N., Eggeling L., Eikmanns B.-J., Galsglat L.,
RA Goessmann A., Hartmann M., Hutmacher K., Kraemer R., Linke B.,
RA McHardy A.C., Meyer F., Moeckel B., Pfeifferle W., Piehler A.,
RA Rey D.A., Rueckert C., Rupp O., Sahn H., Wendisch V.F., Wiegand I.,
RA Tauch A., Tuerker C., Tuerker O., Tuerker H., Wiegand I.,
RT "The complete Corynebacterium glutamicum ATCC 13032 genome sequence
RT and its impact on the production of L-aspartate-derived amino acids
RT and vitamins."
RL J. Biotechnol. 104:5-25(2003).
CC -1- SIMILARITY: Contains 1 PPIase cyclophilin-type domain.
DR EMBL: BX927148; CAP16602.1; -; Genomic_DNA.
DR HSSP: Q27450; 1A58.
DR SMR: Q6M8Y0; 5-173.
DR GO: GO:0016853; P:isomerase activity; IEA.
DR GO: GO:0003755; P:peptidyl-prolyl cis-trans isomerase activity; IEA.
DR GO: GO:0006457; P:protein folding; IEA.
DR InterPro: IPR002130; CSA_PPIase.
DR Pfam: PF00160; Pro_Isoomerase; 1.
DR PRINTS: PR00153; CSAPPISMRASE.
DR PROSITE: PS00170; CSA_PPIASE_1; 1.
DR PROSITE: PS00072; CSA_PPIASE_2; 1.
DR Isomerase; Rotamase.
KW Isomerase; Rotamase.
SQ SEQUENCE 174 AA; 18497 MW; 9064515401ECC2F9 CRC64;

Query Match 74.5%; Score 716; DB 2; Length 174;
Best Local Similarity 79.9%; Pred. No. 1.3e-60;
Matches 135; Conservative 11; Mismatches 23; Indels 0; Gaps 0;

Qy      14 TATATHTNRGDIKALFGNHAPKTVANFVGLAQGTQDSTONASGSGPGPFYDGVFHR 73
      5 TATATHTNRGDIITDLFGNHAPETVANFVGLAQGTQDYSANAGDSSEPFYNGSVFHR 64
Db      134 TPLNRRHTIFGEVIDAESQRVVEAISKTATDGNDRPTDPTVIESITIS 182
      125 TPLNRRHTIFGEVIDAESQKVVDATATATDRYDRPADAVIESVEIT 173

Qy      74 VIOGFMIQGGDPPTGTRGCGPGYKFADEHFPELOFDPFPLAAMNAGPGTNGSQFPTVTK 133
      65 VIDGFMIQGGDPPTGTRGCGPGYTFADHFPELFPDAVYLLAAMNAGPGTNGSQFPTVTP 124
Db      134 TPLNRRHTIFGEVIDAESQRVVEAISKTATDGNDRPTDPTVIESITIS 182
      125 TPLNRRHTIFGEVIDAESQKVVDATATATDRYDRPADAVIESVEIT 173

RESULT 7
Q8NUAS CORGL PRELIMINARY; PRT; 187 AA.
ID Q8NUAS;
AC Q8NUAS;
DT 01-OCT-2002 (T-EMBLrel. 22, Created)
DT 01-OCT-2002 (T-EMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)

```

```

DE      Peptidyl-prolyl cis-trans isomerase (Rotamase)-cyclophilin family
DE      (EC 5.2.1.8).
GN OrderedlocusNames=Cg10034;
OS Corynebacterium glutamicum (Brevibacterium flavum).
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacteriaceae; Corynebacteriaceae; Corynebacterium.
OX NCBI_TaxID=1718;
(1) _SEQUENCE.
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;
RC PubMed=12948625; DOI=10.1016/S0168-1656(03)00154-8;
RA Nakagawa S.;
RT "Complete genomic sequence of Corynebacterium glutamicum ATCC 13032."
RT Submitted (MAR-2002) to the EMBL/Genbank/DBJ databases.
CC -1- SIMILARITY: Contains 1 PPIase cyclophilin-type domain.
DR EMBL: BA000036; BAB97427.1; -; Genomic_DNA.
DR HSSP: P30412; 2RMC.
DR SMR: Q8NUAS; 18-186.
DR GO: GO:0016853; P:isomerase activity; IEA.
DR GO: GO:0003755; P:peptidyl-prolyl cis-trans isomerase activity; IEA.
DR GO: GO:0006457; P:protein folding; IEA.
DR InterPro: IPR002130; CSA_PPIase.
DR Pfam: PF00160; Pro_Isoomerase; 1.
DR PRINTS: PR00153; CSAPPISMRASE.
DR PROSITE: PS00170; CSA_PPIASE_1; 1.
DR PROSITE: PS00072; CSA_PPIASE_2; 1.
DR Complete proteome; Isomerase; Rotamase.
KW Complete proteome; Isomerase; Rotamase.
SQ SEQUENCE 187 AA; 19830 MW; FBCC3757D051787F CRC64;

Query Match 74.5%; Score 716; DB 2; Length 187;
Best Local Similarity 79.9%; Pred. No. 1.4e-60;
Matches 135; Conservative 11; Mismatches 23; Indels 0; Gaps 0;

Qy      14 TATATHTNRGDIKALFGNHAPKTVANFVGLAQGTQDSTONASGSGPGPFYDGVFHR 73
      18 TATATHTNRGDIITDLFGNHAPETVANFVGLAQGTQDYSANAGDSSEPFYNGSVFHR 77
Db      134 TPLNRRHTIFGEVIDAESQRVVEAISKTATDGNDRPTDPTVIESITIS 182
      78 VIDGFMIQGGDPPTGTRGCGPGYTFADHFPELFPDAVYLLAAMNAGPGTNGSQFPTVTP 137
Qy      134 TPLNRRHTIFGEVIDAESQRVVEAISKTATDGNDRPTDPTVIESITIS 182
      138 TPLNRRHTIFGEVIDAESQKVVDATATATDRYDRPADAVIESVEIT 166
Db      138 TPLNRRHTIFGEVIDAESQKVVDATATATDRYDRPADAVIESVEIT 166

RESULT 8
Q6NKJ3 CORGL PRELIMINARY; PRT; 180 AA.
ID Q6NKJ3;
AC Q6NKJ3;
DT 05-JUL-2004 (T-EMBLrel. 27, Created)
DT 05-JUL-2004 (T-EMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (T-EMBLrel. 27, Last annotation update)
DE Problem peptidyl-prolyl cis-trans isomerase A (EC 5.2.1.8).
GN Name:pplA; OrderedlocusNames=DIP0025;
OS Corynebacterium diptheriae.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacteriaceae; Corynebacteriaceae; Corynebacterium.
OX NCBI_TaxID=1717;
(1) _SEQUENCE.
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=Biotype gravis / NCTC 13129;
RC MEDLINE=22965443; PubMed=14602910; DOI=10.1093/nar/9/gk9874;
RC Cerdano-Tarraga A.-M., Estratou A., Dover L.G., Holden M.T.G.,
RA Pallen M.J., Bentley S.D., Besra G.S., Churcher C.M., James K.D.,
RA De Zorja A., Chillingworth T., Cronin A., Dowd L., Felwell T.,
RA Hamlin N., Holroyd S., Jagsels K., Moule S., Quail M.A.,
RA Rabinovitch E., Rutherford K.M., Thomson N.R., Unwin L.,
RA Whitehead S., Barrett B.G., Parkhill J.,
RT "The complete genome sequence and analysis of Corynebacterium
RT diptheriae NCTC13129."
RL Nucleic Acids Res. 31:6516-6523(2003).
CC -1- SIMILARITY: Contains 1 PPIase cyclophilin-type domain.
DR EMBL: BX248354; CAB48536.1; -; Genomic_DNA.

```

DR HSSP; Q27450; 1A33.
 DR SMR; Q6NKC3; 9-179.
 DR GO; GO:0016853; F:isomerase activity; IEA.
 DR GO; GO:0003755; P:peptidyl-prolyl cis-trans isomerase activity; IEA.
 DR GO; GO:0006457; P:protein folding; IEA.
 DR InterPro; IPR002130; CSA_PPIase.
 DR InterPro; IPR002135; Pro_1in_serpin.
 DR Pfam; PF00160; Pro_Isomerase; 1.
 DR PRINTS; PS00153; CSAPISMASE.
 DR PROSITE; PS00170; CSA_PPIASE_1; 1.
 DR PROSITE; PS50072; CSA_PPIASE_2; 1.
 DR PROSITE; PS00284; SERPIN; UNKNOWN_1.
 DR Complete proteome; Isomerase; Rotamase.
 DR SEQUENCE 180 AA; 19442 MW; F0468373BDDF598 CRC64;

Query Match 73.9%; Score 710; DB 2; Length 180;
 Best Local Similarity 78.4%; Pred. No. 5e-60;
 Matches 134; Conservative 11; Mismatches 26; Indels 0; Gaps 0;

QY 12 LATATATLHTNRGDIKIALFGNHAKPTVANFVGLAQSTKDYSTONASGSPGSPFYDGAVF 71
 DT 25-OCT-2004 (TrEMBLrel. 28, Created)
 DB 9 LKATATLHTNRGDVSELFGNHAKPTVENFVTLANGTAYKTENASGTNEGPFYDGAVF 68
 QY 72 HRYIOGFMIOGSPPTGTGRGPGYKFPADBFHPELOFDPKPYLLAMNAGPQTNGSQFFITV 131
 DB 69 HRVYIDFMIQGGDPPTGTGRGPGYKFPADBFHPELOFDPKPYLLAMNAGPQTNGSQFFITV 128
 QY 132 GKTPHLNRHTIFGEVIDASORVVEAISKATDGNDRPTDPVVISITIS 182
 DB 129 VPTPLNRHTIFGEVIDASORVVEAISKATDGNDRPTDPVVISITIS 179

RESULT 9
 Q8FUI9 CORP PRELIMINARY; PRT; 202 AA.

AC Q8FUI9;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DE 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Putative peptidyl-prolyl isomerase.
 OS OrderedLocustNames=CE0029;
 OS Corynebacterium efficiens.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacteriaceae; Corynebacteriaceae; Corynebacterium.
 NCBI_TaxID=152794;
 RN (1)
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=IS-314 / AJ 12310 / DSM 44549 / JCM 11189;
 RX MEDLINE=22723752; PubMed=12840036; DOI=10.1101/gr.1285603;
 RA Nishio Y., Nakamura Y., Kawarabayashi Y., Usuda Y., Kimura E.,
 RA Sugimoto S., Matsui K., Yamagishi A., Kikuchi H., Ikeo K.,
 RA Gojobori T.;
 RT "Comparative complete genome sequence analysis of the amino acid
 RT replacements responsible for the thermostability of Corynebacterium
 RT efficiens.";
 RL Genome Res. 13:1572-1579 (2003).
 CC -1- SIMILARITY: Contains 1 PPIase cyclophilin-type domain.
 DR EMBL; BA000035; BAC16839.1; -, Genomic_DNA.
 DR HSSP; P30412; 2RMC.
 DR SMR; Q8FUI9; 31-200.
 DR GO; GO:0016853; F:isomerase activity; IEA.
 DR GO; GO:0003755; P:peptidyl-prolyl cis-trans isomerase activity; IEA.
 DR GO; GO:0006457; P:protein folding; IEA.
 DR InterPro; IPR002130; CSA_PPIase.
 DR Pfam; PF00160; Pro_Isomerase; 1.
 DR PRINTS; PS00153; CSAPISMASE.
 DR PROSITE; PS00170; CSA_PPIASE_1; 1.
 DR PROSITE; PS50072; CSA_PPIASE_2; 1.
 DR Complete proteome; Isomerase; Rotamase.
 DR SEQUENCE 202 AA; 21829 MW; 814495B2B7C3A818 CRC64;

Query Match 73.7%; Score 708; DB 2; Length 202;
 Best Local Similarity 76.0%; Pred. No. 9e-60;

Matches 133; Conservative 11; Mismatches 31; Indels 0; Gaps 0;
 QY 7 VINSPLATATATLHTNRGDIKIALFGNHAKPTVANFVGLAQSTKDYSTONASGSPGSPFY 66
 DB 26 VVSMTLKATATLHTNRGDISIDILFGNHAKPTVANFVGLAQSTKDYSTONASGSPGSPFY 85
 QY 67 DGAVFHRVYIOGFMIOGSPPTGTGRGPGYKFPADBFHPELOFDPKPYLLAMNAGPQTNGSQ 126
 DB 86 DGAVFHRVIDRFMIQGGDPPTGTGRGPGYKFPADBFHPELOFDPKPYLLAMNAGPQTNGSQ 145
 QY 127 FFIITGKTPHLNRHTIFGEVIDASORVVEAISKATDGNDRPTDPVVISITIS 181
 DB 146 FFIITGKTPHLNRHTIFGEVIDASORVVEAISKATDGNDRPTDPVVISITIS 200

RESULT 10
 Q6A9U1 PROAC PRELIMINARY; PRT; 170 AA.

AC Q6A9U1;
 DT 25-OCT-2004 (TrEMBLrel. 28, Created)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
 DE 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
 DE Probable peptidyl-prolyl cis-trans isomerase A (EC 5.2.1.8).
 OS OrderedLocustNames=PPA0718;
 OS Propionibacterium acnes.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Propionibacteriaceae; Propionibacteriaceae; Propionibacterium.
 NCBI_TaxID=1747;
 RN (1)
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=KFA171202 / DSM 16379;
 RX PubMed=15286373; DOI=10.1126/science.1100330;
 RA Brueggemann H., Henne A., Hoerster F., Liesegang H., Wenzel A.,
 RA Strittmatter A., Hujer S., Duerre P., Gottschalk G.;
 RT "The complete genome sequence of Propionibacterium acnes, a commensal
 RT of human skin.";
 RL Science 305:671-673 (2004).
 CC -1- SIMILARITY: Contains 1 PPIase cyclophilin-type domain.
 DR EMBL; AE017283; AAT82475.1; -, Genomic_DNA.
 DR GO; GO:0016853; F:isomerase activity; IEA.
 DR GO; GO:0003755; P:peptidyl-prolyl cis-trans isomerase activity; IEA.
 DR GO; GO:0006457; P:protein folding; IEA.
 DR InterPro; IPR002130; CSA_PPIase.
 DR Pfam; PF00160; Pro_Isomerase; 1.
 DR PRINTS; PS00153; CSAPISMASE.
 DR PROSITE; PS00170; CSA_PPIASE_1; 1.
 DR PROSITE; PS50072; CSA_PPIASE_2; 1.
 DR Complete proteome; Isomerase; Rotamase.
 DR SEQUENCE 170 AA; 18378 MW; EADPBLA83BE6546 CRC64;

Query Match 63.6%; Score 611; DB 2; Length 170;
 Best Local Similarity 69.0%; Pred. No. 1.6e-50;
 Matches 116; Conservative 15; Mismatches 37; Indels 0; Gaps 0;

QY 15 ATATLHTNRGDIKIALFGNHAKPTVANFVGLAQSTKDYSTONASGSPGSPFYDGAVFHRV 74
 DB 3 STATLHTNRGDIVNLFADQAPKTVDNFGVLAGSTKEVDPRHGTGKFGDGLFHRV 62
 QY 75 IOGFMIOGSPPTGTGRGPGYKFPADBFHPELOFDPKPYLLAMNAGPQTNGSQFFITV 134
 DB 63 IOGFMIOGSPPTGTGRGPGYKFPADBFHPELOFDPKPYLLAMNAGPQTNGSQFFITV 122
 QY 135 PHLRHTIFGEVIDASORVVEAISKATDGNDRPTDPVVISITIS 182
 DB 123 PHLRHTIFGEVIDASORVVEAISKATDGNDRPTDPVVISITIS 170

RESULT 11
 Q82FC6 STRAW PRELIMINARY; PRT; 186 AA.

AC Q82FC6;
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)


```

DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE Putative peptidyl-prolyl cis-trans isomerase.
GN Name=pp1A; OrderedLocNames=SAV4329;
OS Streptomyces avermitilis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomyces; Streptomycetaceae; Streptomyces.
OX NCBI_TaxId=33903;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=NA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX MEDLINE=21477403; PubMed=11572948; DOI=10.1073/pnas.211433198;
RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osomoe T.,
RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
RA "Genome sequence of an industrial microorganism Streptomyces
RT avermitilis: deducing the ability of producing secondary
RT metabolites.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=NA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX MEDLINE=22608306; PubMed=12692562; DOI=10.1038/nbt820;
RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
RA Sakaki Y., Hattori M., Omura S.;
RA "Complete genome sequence and comparative analysis of the industrial
RT microorganism Streptomyces avermitilis.";
RL Nat. Biotechnol. 21:526-531(2003).
CC -1- SIMILARITY: Contains 1 PPIase cyclophilin-type domain.
DR EMBL: BAO00030; BAC72041.1; -; Genomic DNA.
DR HSSP: Q27450; 1A58.
DR GO: GO:0016853; P:isomerase activity; IEA.
DR GO: GO:0003755; P:peptidyl-prolyl cis-trans isomerase activity; IEA.
DR GO: GO:0006457; P:protein folding; IEA.
DR InterPro: IPR002130; CSA_PPIase.
DR Pfam: PF00160; Pro_Isomerase; 1.
DR PRINTS: PR00153; CSAPISMRASE.
DR PROSITE: PS00170; CSA_PPIASE_1; 1.
DR PROSITE: PS50072; CSA_PPIASE_2; 1.
DR Complete proteome; Isomerase; Rotamase.
SQ SEQUENCE 186 AA; 20307 MW; 80BB7C577A5E05 CRC64;

Query Match 56.6%; Score 543.5; DB 2; Length 186;
Best Local Similarity 65.3%; Pred. No. 5,4e-44;
Matches 109; Conservative 17; Mismatches 38; Indels 3; Gaps 3;

QY 17 ATLHTNRGDIKIALFGNHPKVTANFVGLAQGTGYSTONASGSGPS-GPEYDGVFPRVY 75
DB 7 ATLKTSQGDIAVALLFNHHPKVTANFVGLATGEREM-TNPATSEKSKDKLYDGVFPRVY 65
QY 76 QGFMIQGGDPPTGTGRGPGYKPADBFHPBLQDPKPYLLMANNGPQTNGSQPFITVGKTP 135
DB 66 SGFMIGGGDPPLGNGTDDPGYQFQDEBHPDLRFKPYLLMANNGPQTNGSQPFITVAPTA 125
QY 136 HUNRRRTTIGEVDASQRYVEAISKTAID-GNDRPTDPPVISTIT 181
DB 126 WNRKRTTIGEYVDVASQKVDAIAGRTNPGTKRPVSDIVISVVI 172

RESULT 12
OGAHNS_LEIIX PRELIMINARY; PRT; 179 AA.
ID OGAHNS_LEIIX
AC OGAHNS;
DT 25-OCT-2004 (Tremblrel. 28, Created)
DT 25-OCT-2004 (Tremblrel. 28, Last sequence update)
DT 25-OCT-2004 (Tremblrel. 28, Last annotation update)
DE Peptidyl-prolyl cis-trans isomerase.
GN Name=pp1B; OrderedLocNames=Lxx00160;
GN Leifsonia xyli (subsp. xyli).
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Micrococciaceae; Microbacteriaceae; Leifsonia.
OX NCBI_TaxId=59736;
RN [1]
RP NUCLEOTIDE SEQUENCE.

```

```

RC STRAIN=CTCB07;
RX PubMed=15305603;
RA Monteiro-Vitorello C.B., Camargo L.E.A., Van Sluys M.A.,
RA Kitajima J.P., Truffi D., do Amaral A.M., Haraoka R.,
RA de Oliveira J.C.F., Wood D., de Oliveira M.C., Miyaki C.Y.,
RA Takita M.A., da Silva A.C.R., Furlan L.R., Carraro D.M., Camarotte G.,
RA Almeida N.F., Jr., Carrier H., Coutinho L.L., El-Dorri H.A.,
RA Ferro M.I.T., Gagliardi P.R., Gigliotti E., Goldman M.H.S.,
RA Goldman G.H., Kimura E.T., Fairo B.S., Kuraue E.E., Lemos E.G.M.,
RA Lemos M.V.F., Mauro S.M.Z., Machado M.A., Martino C.L., Menck C.F.,
RA Nunes L.R., Oliveira R.C., Pereira G.G., Siqueira W., de Souza A.A.,
RA Trai S.M., Zanca A.S., Simpson A.U.G., Brumley S.M., Seudal J.C.;
RA "The genome sequence of the Gram-positive sugarcane pathogen Leifsonia
RT xyli subsp. xyli.";
RL Mol. Plant Microbe Interact. 17:827-836(2004).
CC -1- SIMILARITY: Contains 1 PPIase cyclophilin-type domain.
DR EMBL: AE016822; AAT88120.1; -; Genomic DNA.
DR GO: GO:0016853; P:isomerase activity; IEA.
DR GO: GO:0003755; P:peptidyl-prolyl cis-trans isomerase activity; IEA.
DR GO: GO:0006457; P:protein folding; IEA.
DR InterPro: IPR002130; CSA_PPIase.
DR Pfam: PF00160; Pro_Isomerase; 1.
DR PRINTS: PR00153; CSAPISMRASE.
DR PROSITE: PS00170; CSA_PPIASE_1; 1.
DR PROSITE: PS50072; CSA_PPIASE_2; 1.
DR Complete proteome; Isomerase; Rotamase.
SQ SEQUENCE 179 AA; 19233 MW; 4204C1783128784F CRC64;

Query Match 54.9%; Score 528; DB 2; Length 179;
Best Local Similarity 60.7%; Pred. No. 1,6e-42;
Matches 105; Conservative 21; Mismatches 41; Indels 6; Gaps 3;

QY 14 TATATLHTNRGDIKIALFGNHPKVTANFVGLAQGTGYSTONASGSGPSG-PFYDGVF 72
DB 5 TAVATLTYTNGDKLVNLFNHPKVTANFVGLATGEREM-TNPATSEKNTPLDGVF 63
QY 73 RVIQGFMIQGGDPPTGTGRGPGYKPADBFHPBLQDPKPYLLMANNGP----GTNGSQPF 128
DB 64 RIIPGFMIQGGDPICGGIGGPGYQFQDEINPELDFAQPYLLMANNGSPGSGGTNGSQPF 123
QY 129 ITVGTGPHLNRRTTIGEYVDASQRYVEAISKTAIDGNDRPTDPPVISTIT 181
DB 124 ITVGTPTWLGQKHSIFGEVDADSRKVDKLABPTDAGDRPLDVIISVVI 176

RESULT 13
PPIB_STRCH
ID PPIB_STRCH STANDARD; PRT; 175 AA.
AC P77949;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Peptidyl-prolyl cis-trans isomerase B (EC 5.2.1.8) (PPIase B)
DE (Rotamase B) (Cyclophilin scCyPB) (S-cyclophilin).
GN Name=cyPB;
OS Streptomyces chrysomallus.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomyces; Streptomycetaceae; Streptomyces.
OX NCBI_TaxId=1892;
RN [1]
RP NUCLEOTIDE SEQUENCE [GENOMIC DNA], AND CHARACTERIZATION.
RX MEDLINE=9717779; PubMed=9025285;
RA Pahl A., Gewies A., Keller U.;
RA "scCyPB is a novel second cytosolic cyclophilin from Streptomyces
RT chrysomallus which is phylogenetically distant from scCyPB.";
RL Microbiology 143:117-126(1997).
CC -1- FUNCTION: PPIase accelerate the folding of proteins. It catalyzes
CC the cis-trans isomerization of proline imidic peptide bonds in
CC oligopeptides.
CC -1- CATALYTIC ACTIVITY: Peptidylproline (omega=180) = peptidylproline
CC (omega=0).
CC -1- ENZYME REGULATION: Inhibited by cyclosporin A (Csa).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.

```


Db	4	IATAKATIHSTLGDIVVNLFGNHA	PVTYKNFVGLATGEQAMTHPETGEDTKTGPLYNGTI	63
Qy	71	FHRVIOGFMIQSGDPTGTGRGSPGYKFADE	FHPELOFDKPYLLAMANA----	GPGTNGSQ 126
Db	64	FHRITIDFMIQAGDPLGRGVGSPGYKFDDEI	HPLETFNEBPYKLANANAGIQMGKGTNGSQ	123
Qy	127	PFITVGTKTPHLNRRTIFGEVIDAESORVVEA	ISKTATDGNDRPTDPVYESITI	181
Db	124	PFITTIPTDWLQGHKSIFGEVADDESKCVDA	LEGVRTGMGDRPVBDVINSIDV	178

Search completed: April 14, 2006, 17:32:25
 Job time : 70.152 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: April 14, 2006, 17:32:59 ; Search time 16.829 Seconds

(without alignments)
894.107 Million cell updates/sec

Title: US-10-620-246-12

Perfect score: 961

Sequence: 1 MACDSVTNSPLATATATLH.....ATDGNRPDPVIESITIS 182

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA:*

1: /cgn2_6/ptodata/1/1aa/5 COMB.pep:*

2: /cgn2_6/ptodata/1/1aa/6 COMB.pep:*

3: /cgn2_6/ptodata/1/1aa/H COMB.pep:*

4: /cgn2_6/ptodata/1/1aa/PCUS COMB.pep:*

5: /cgn2_6/ptodata/1/1aa/R3 COMB.pep:*

6: /cgn2_6/ptodata/1/1aa/backfill1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	961	100.0	182	2	US-09-050-739-12
2	494.5	51.5	182	2	US-09-902-540-13998
3	403.5	42.0	182	2	US-09-583-110-1345
4	403.5	42.0	182	2	US-09-107-433-4470
5	359	37.4	186	2	US-09-270-767-13856
6	359	37.4	186	2	US-09-270-767-49073
7	347.5	36.2	161	2	US-09-136-442-1
8	347.5	36.2	161	2	US-09-440-828-1
9	343.5	35.7	203	2	US-09-134-001C-3111
10	334	34.8	175	2	US-09-134-000C-3739
11	333.5	34.7	472	2	US-09-088-425-1
12	333.5	34.7	472	2	US-09-639-378A-1
13	333.5	34.7	472	2	US-09-999-833A-245
14	333.5	34.7	472	2	US-10-020-445A-245
15	323	33.6	123	1	US-08-482-728A-8
16	322.5	33.6	282	2	US-09-136-442-4
17	322.5	33.6	282	2	US-09-440-828-4
18	322.5	33.6	520	1	US-08-482-728A-4
19	322.5	33.6	520	2	US-09-028-366-5
20	322.5	33.6	520	2	US-09-715-285-5
21	319.5	33.2	169	1	US-08-145-995A-7
22	319.5	33.2	169	1	US-08-451-747-7
23	319.5	33.2	169	2	US-09-134-852-7
24	315	32.8	212	1	US-08-142-897-5
25	315	32.8	212	1	US-08-248-796A-19586
26	310.5	32.3	212	2	US-09-538-092-1126
27	309.5	32.2	238	2	US-09-949-002-510

28	309	32.2	203	2	US-10-043-142-10	Sequence 10, Appl
29	309	32.2	203	2	US-09-806-399-10	Sequence 10, Appl
30	305.5	31.8	224	2	US-09-134-000C-6764	Sequence 6764, Appl
31	300	31.2	212	2	US-10-043-142-5	Sequence 5, Appl
32	300	31.2	212	2	US-09-806-399-5	Sequence 5, Appl
33	295	30.7	147	2	US-09-136-442-3	Sequence 3, Appl
34	295	30.7	147	2	US-09-440-828-3	Sequence 3, Appl
35	291.5	30.3	207	2	US-10-043-142-11	Sequence 11, Appl
36	291.5	30.3	207	2	US-09-806-399-11	Sequence 11, Appl
37	291	30.3	208	1	US-08-142-897-7	Sequence 7, Appl
38	291	30.3	208	2	US-10-043-142-12	Sequence 12, Appl
39	291	30.3	208	2	US-09-806-399-12	Sequence 12, Appl
40	291	30.3	208	2	US-09-538-092-994	Sequence 994, App
41	287.5	29.9	176	1	US-08-145-995A-3	Sequence 3, Appl
42	287.5	29.9	176	1	US-08-451-747-3	Sequence 3, Appl
43	287.5	29.9	176	2	US-09-134-852-3	Sequence 3, Appl
44	287.5	29.9	269	2	US-09-028-366-6	Sequence 6, Appl
45	287.5	29.9	269	2	US-09-715-285-5	Sequence 6, Appl

ALIGNMENTS

RESULT 1									
US-09-050-739-12									
Sequence 12, Application US/09050739									
Patent No. 6641814									
GENERAL INFORMATION:									
APPLICANT: ANDERSEN, Peter									
APPLICANT: NIELSEN, Rikke									
APPLICANT: OETTINGER, Thomas									
APPLICANT: RASMUSSEN, Peter Blak									
APPLICANT: ROSENKRANDS, Ida									
APPLICANT: WELDRING, Karin									
APPLICANT: FLORIO, Walter									
TITLE OF INVENTION: NUCLEIC ACIDS FRAGMENTS AND POLYPEPTIDE FRAGMENTS									
FILE REFERENCE: 670001-2002.1									
CURRENT APPLICATION NUMBER: US/09/050, 739									
CURRENT FILING DATE: 1998-03-30									
EARLIER APPLICATION NUMBER: 0376/97									
EARLIER FILING DATE: 1997-04-02									
EARLIER APPLICATION NUMBER: 1277/97									
EARLIER FILING DATE: 1997-11-10									
EARLIER APPLICATION NUMBER: 60/044, 624									
EARLIER FILING DATE: 1997-04-18									
EARLIER APPLICATION NUMBER: 60/070, 488									
EARLIER FILING DATE: 1998-01-05									
NUMBER OF SEQ ID NOS: 173									
SOFTWARE: Patentin Ver. 2.0									
SEQ ID NO 12									
LENGTH: 182									
TYPE: PRT									
ORGANISM: Mycobacterium tuberculosis									
US-09-050-739-12									
Query Match									
Best Local Similarity 100.0%; Score 961; DB 2; Length 182;									
Matches 182; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY	1	MACDSVTNSPLATATATLTNRGDIKIALFGNHAPKTVANFYGLAQGTQSYTONASGG	60						
DB	1	MACDSVTNSPLATATATLTNRGDIKIALFGNHAPKTVANFYGLAQGTQSYTONASGG	60						
QY	61	PSGPFYDGAFFHVIQGMIOGSDPTGTGSGGQYKFADEFHPELOPKYLLAMNAGP	120						
DB	61	PSGPFYDGAFFHVIQGMIOGSDPTGTGSGGQYKFADEFHPELOPKYLLAMNAGP	120						
QY	121	GTGSGPFYVGTGTPHNRHTTIFGVYDAESQRYVBAISKATDGNDRPTDPVIESIT	180						
DB	121	GTGSGPFYVGTGTPHNRHTTIFGVYDAESQRYVBAISKATDGNDRPTDPVIESIT	180						
QY	181	IS 182							

Db 181 IS 182

RESULT 2

US-09-902-540-13998
; Sequence 13998, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; NUMBER OF SEQ ID NOS: 16825
; LENGTH: 182
; TYPE: PRF
; ORGANISM: Myxococcus xanthus
US-09-902-540-13998

Query Match 51.5%; Score 494.5; DB 2; Length 182;

Best Local Similarity 60.5%; Pred. No. 1,3e-46;

Matches 101; Conservative 12; Mismatches 51; Indels 3; Gaps 2;

QY 17 ATLHNRGDIKIALFGNAPKTVANFVGLAOGTKDYSTONASGGPSG-PYDGAVERHVI 75

Db 9 ATLETQGAIVLRFSDAPKTVANFVGLAOGTKDYSTONASGGPSG-PYDGAVERHVI 68

QY 76 QGEMIOGGDPTGGRGGRGKFADEFEHPELOPKPYLLANANAGPGNSQPFITVCKTP 135

Db 69 PGFMIOGGDPTGGRGGRGKFADEFEHPELOPKPYLLANANAGPGNSQPFITVCKTP 128

QY 136 HLNRHRTIFGEVDAESQRYEVAISKATDGNDRPTDPTVIESITIS 182

Db 129 YLNRHRTIFGEVDAESQRYEVAISKATDGNDRPTDPTVIESITIS 173

RESULT 3

US-09-583-110-3345
; Sequence 3345, Application US/09583110
; Patent No. 6699703
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al.
; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
; FILE REFERENCE: PAT00-07A
; CURRENT APPLICATION NUMBER: US/09/583,110
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/107,433
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/085,131
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: US 60/051,553
; PRIOR FILING DATE: 1997-07-02
; NUMBER OF SEQ ID NOS: 5322
; SEQ ID NO 3345
; LENGTH: 466
; TYPE: PRF
; ORGANISM: Streptococcus pneumoniae
US-09-583-110-3345

Query Match 42.0%; Score 403.5; DB 2; Length 466;

Best Local Similarity 44.5%; Pred. No. 5.1e-36;

Matches 93; Conservative 23; Mismatches 42; Indels 51; Gaps 7;

QY 3 DCDSTVNSPLATATATLHNRGDIKIALFGNAPKTVANFVGLAOGTKDYSTONASGGPS 62

Db 277 DIETV-EGPL---ATIKTNHGDRLIKLPENAPKTVANFVSL---SKD----- 317

QY 63 GPEYDGAVERHVIQGMIOGSDPTGGRGP--GYKFADEFEHPELOPKPYLLANANAG 119

Db 318 -GYDGVIFHRIIKDPMIOGSDPTGGMGBES1YGSFDEFESEL-YNIRGLSNANAG 375

QY 120 PGTNSQPFIT-----TVGKTPHLNRHRTIFGEVDAES 152

Db 376 PNTNSQPFITVQNHLPYSKKEITRGWPEPIAEIYANOGCTPHLDRHRTVFGQLADEAS 435

QY 153 QRVEAISKATDGNDRPTDPTVIESITI 181

Db 436 YAVLDVIAVETGAMKPDVDVITETI 464

RESULT 4

US-09-107-433-4470

; Sequence 4470, Application US/09107433

; Patent No. 6800744

; GENERAL INFORMATION:

; APPLICANT: Lynn A Doucette-Stamm and David Bush

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID

; SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE

; THERAPEUTICS

; NUMBER OF SEQUENCES: 5206

; CORRESPONDENCE ADDRESS:

; ADDRESSER: GENOME THERAPEUTICS CORPORATION

; STREET: 100 Beaver Street

; CITY: Waltham

; STATE: Massachusetts

; COUNTRY: USA

; ZIP: 02354

; COMPUTER READABLE FORM:

; MEDIUM TYPE: CD-ROM ISO9660

; OPERATING SYSTEM: <Unknown>

; SOFTWARE: <Unknown>

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/107,433

; FILING DATE: 30-Jun-1998

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 60/085131

; FILING DATE: May 12, 1998

; APPLICATION NUMBER: 60/051553

; FILING DATE: July 2, 1997

; ATTORNEY/AGENT INFORMATION:

; NAME: Arinello, Pamela Deneke

; REGISTRATION NUMBER: 40,489

; REFERENCE/DOCKET NUMBER: GTC-011

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (781)893-5007

; TELEFAX: (781)893-8277

; INFORMATION FOR SEQ ID NO: 4470:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 472 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; HYPOTHEICAL: YES

; ORIGINAL SOURCE:

; ORGANISM: Streptococcus pneumoniae

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: (B) LOCATION 1...472

; SEQUENCE DESCRIPTION: SEQ ID NO: 4470:

Query Match 42.0%; Score 403.5; DB 2; Length 472;

Best Local Similarity 44.5%; Pred. No. 5.2e-36;

Matches 93; Conservative 23; Mismatches 42; Indels 51; Gaps 7;

QY 3 DCDSTVNSPLATATATLHNRGDIKIALFGNAPKTVANFVGLAOGTKDYSTONASGGPS 62

```

Db      283  DIEAT-EGSL-----ATIKNHDGLRIKLPPEHAKPTVANFVSLD---SKD----- 323
Oy      63  GPFGCAVAFHRYVIOGFMIQGGDPYTGRCGP---GYKPADDEFHPELQFPDKPYLLMANANG 119
      :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db      324  -GYVGAVFPHRIILIKOFMIQGGDPYTGRCGSESIYGBSFDEPSEEL--YNIKGLSMVANG 381
      :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Oy      120  PGTNGSOPEPI-----TYGKPTPLNRRTITFGEVIDAAS 152
      |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db      382  PNTNGSOFFIVONOHLPYSKKEITTRGGMWEPPLAEIYANGGTFPLDLRRITFVGLDAEAS 441
      :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Oy      153  ORVVEAISKTATDGNDRPTDPPVVISITTI 181
      :||:||||:||||:||||:||||:||||:||||:||||:||||:
Db      442  YAVLDVIAVETGAMDKPYEDVVIETITEL 470

RESULT 5
US-09-270-767-33856
: Sequence 33856, Application US/09270767
: Patent No. 6703491
: GENERAL INFORMATION:
: APPLICANT: Homburger et al.
: TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
: FILE REFERENCE: File Reference: 7326-094
: CURRENT APPLICATION NUMBER: US/09/270,767
: CURRENT FILING DATE: 1999-03-17
: NUMBER OF SEQ ID NOS: 62517
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 33856
: LENGTH: 186
: TYPE: PRT
: ORGANISM: Drosophila melanogaster
US-09-270-767-33856

```

	Query Match	37.4%	Score 359;	DB 2;	Length 186;	
	Best Local Similarity	46.9%	Pred. No. 1.2e-31;			
	Matches	76;	Conservative	21;	Mismatches	45; Indels 20; Gaps 3;
Oy	18 TLATNRGDIKIALPFGNHAKPTVANFVGLAOGTKDYSTONASGSGSPGYDGAVFRRVIQG	77	:	:	:	:
Dd	34 TLETSMBEITVELYKHAHPNTRCRNPALSR-----RGYYNNVVFFRIIRD	78	:	:	:	:
Oy	78 FMIOGGDPGTGRGQP---GYKPADEFPELQFDKPILLANMAGPGTNSQPIIVGKT	134	:	:	:	:
Dd	79 FMIQGGDPGTGRGGSAYGSEFPADELHGLRHTGAILSMANSQPDTPNGSQPFITLAFT	138	:	:	:	:
Oy	135 PHLARRHRTIRGEVIDAESQRVVEALSKAATYGNDPRDTPPVYI	176	:	:	:	:
Dd	139 QWLDGKHITIFGRVY--TGMEVVKKRIGWETDKNDRPFLRLI	178	:	:	:	:
	RESULT 6					
	US-09-270-767-49073					
	/ Sequence 49073, Application US/09270767					
	/ Patent NO. 6703491					
	/ GENERAL INFORMATION:					
	/ APPLICANT: Homburger et al.					
	/ TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster					
	/ FILE REFERENCE: File Reference: 7356-094					
	/ CURRENT APPLICATION NUMBER: US/09/270,767					
	/ CURRENT FILING DATE: 1999-03-17					
	/ NUMBER OF SEQ ID NOS: 62517					
	/ SOFTWARE: PatentIn Ver. 2.0					
	/ SEQ ID NO 49073					
	/ LENGTH: 186					
	/ TYPE: PRT					
	/ ORGANISM: Drosophila melanogaster					
	US-09-270-767-49073					
	Query Match	37.4%	Score 359;	DB 2;	Length 186;	
	Best Local Similarity	46.9%	Pred. No. 1.2e-31;			
	Matches	76;	Conservative	21;	Mismatches	45; Indels 20; Gaps 3;
Oy	18 TLATNRGDIKIALPFGNHAKPTVANFVGLAOGTKDYSTONASGSGSPGYDGAVFRRVIQG	77	:	:	:	:

[illegible]

```

RESULT 7
US-09-136-442-1
/ Sequence 1, Application US/09136442
/ Patent No. 6030825
/ GENERAL INFORMATION:
/ APPLICANT: Hillman, Jennifer L.
/ APPLICANT: Corley, Neil C.
/ APPLICANT: Guegler, Karl J.
/ APPLICANT: Patterson, Chandra
/ TITLE OF INVENTION: CYCLOPHILIN-TYPE PEPTIDYL-PROLYL CIS/TRANS ISOMERASE
/ FILE REFERENCE: PF-0582 US
/ CURRENT APPLICATION NUMBER: US/09/136,442
/ CURRENT FILING DATE: 1998-08-19
/ NUMBER OF SEQ. ID NOS: 4
/ SOFTWARE: PERL Program
/ SEQ ID NO 1
/ LENGTH: 161
/ TYPE: PR1
/ ORGANISM: Homo sapiens
/ FEATURE: -
/ OTHER INFORMATION: 2925455
/ US-09-136-442-1

```

	Query Match	36.2%	Score 347.5;	DB 2;	Length 161;
	Best Local Similarity	44.4%;	Pred. No. 1.8e-30;		
	Matches	75;	Conservative	26;	Mismatches 49; Indels 19; Gaps 3.
QY	16 TATLHTNRGDIKIALFGNHA	PKTVANFVGLAQGTIDYSTONASGSGPSGFYDGAVFHRI	75		
	:	:::::	:::::	:	: :::: ::
Db	2 SVYLTHDVGDIKIEVFCE	RTPKTCENPILALC-----	-ASNYYNGCIPHRNI	46	
QY	76 QGFMIQGSDPTGCRGGP---	GKYFADEFHELOFDKPYLLAMANAAGPGTNGSQFTITVG	132		
	:	:::::	:::::	:	: :::: ::
Db	47 KGFNVQTGDPETGCRGNIS	IWGKKFEDEYSBYLLKHNRGVAVSMANNPNNTGSOFTITYG	106		
QY	133 KTHPLNRRHTI	FGEVIDAASORVVEALS	KATADGNDRPTDVPVIESITI	181	
	:	:::::	:::::	:	: :::: ::
Db	107 KQPLMDKTYVFGKAYIDG-	LSTIDELEKL	LPVNERTYAPLNDVAHKIDITI	154	

```

RESULT 8
US-09-440-828-1
; Sequence 1, Application US/09440828
; Patent No. 6458575
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Corley, Neil C.
; APPLICANT: Guegler, Karl J.
; APPLICANT: Patterson, Chandra
; TITLE OF INVENTION: CYCLOPHILIN-TYPE PEPTIDYL-PROLYL CIS/TRANS ISOMERASE
; FILE REFERENCE: PF-0582 US
; CURRENT APPLICATION NUMBER: US/09/440,828
; CURRENT FILING DATE: 1999-11-15
; PRIOR APPLICATION NUMBER: 09/136,442
; PRIOR FILING DATE: 1998-08-19
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PERL Program
; SEQ ID NO 1
; LENGTH: 161
; TYPE: PRT
; ORGANISM: Homo sapiens

```


SEQUENCE CHARACTERISTICS:
LENGTH: 472 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: BRAINON01
CLONE: 2291164
US-09-088-425-1

Query Match 34.7%; Score 333.5; DB 2; Length 472;
Best Local Similarity 40.8%; Pred. No. 2.8e-26;

Matches 71; Conservative 24; Mismatches 60; Indels 19; Gaps 3;

Db 11 PLATATATLTNRGDIKIALFGNHAPEKTVANFVGLAQGTQDYSTQNASGSGSPGYDCAV 70
9 PPTNGKVLTKTTADIDIELMSKEAPKACRPFQLCEA-----YDNTI 53
QY 71 FHRVIOGFMIQGGDPPTGTGRGP--GYKPADBFPHLOPKPYLLAMANNAGPPTNGSQF 127
54 FHRVVPGFIVQGGDPPTGTSGGSBSIYGAPFQDEPHSRIRFRRRGLVAMANNAGSHDNGSQF 113
QY 128 FIVTGTKPHLNRRHTIGEVIDAESQVVEAISKATDGNDRPTDPVVISITI 181
114 FPTLGRADBLNNKRTIFGKVTGDTVYNNLR-LISEVDIDDDERPHNPHKITSCEV 166
Db

RESULT 12

US-09-639-378A-1
Sequence 1, Application US/09639378A

Patent No. 6524838
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Lal, Preeti
APPLICANT: Corley, Neil C.
APPLICANT: Patterson, Chandra
APPLICANT: Baughn, Mariah R.
TITLE OF INVENTION: HUMAN PEPTIDYL-PROLYL ISOMERASASES
FILE REFERENCE: PE-0529-1 DIV
CURRENT APPLICATION NUMBER: US/09/639,378A
CURRENT FILING DATE: 2000-08-15
PRIOR APPLICATION NUMBER: US 09/088,425
PRIOR FILING DATE: 1998-06-01
NUMBER OF SEQ ID NOS: 4
SOFTWARE: PERL Program
SEQ ID NO 1
LENGTH: 472
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No. 6524838 2291164CD1
US-09-639-378A-1

Query Match 34.7%; Score 333.5; DB 2; Length 472;
Best Local Similarity 40.8%; Pred. No. 2.8e-26;

Matches 71; Conservative 24; Mismatches 60; Indels 19; Gaps 3;

QY 11 PLATATATLTNRGDIKIALFGNHAPEKTVANFVGLAQGTQDYSTQNASGSGSPGYDCAV 70
9 PPTNGKVLTKTTADIDIELMSKEAPKACRPFQLCEA-----YDNTI 53
Db 71 FHRVIOGFMIQGGDPPTGTGRGP--GYKPADBFPHLOPKPYLLAMANNAGPPTNGSQF 127
54 FHRVVPGFIVQGGDPPTGTSGGSBSIYGAPFQDEPHSRIRFRRRGLVAMANNAGSHDNGSQF 113
QY 128 FIVTGTKPHLNRRHTIGEVIDAESQVVEAISKATDGNDRPTDPVVISITI 181
114 FPTLGRADBLNNKRTIFGKVTGDTVYNNLR-LISEVDIDDDERPHNPHKITSCEV 166
Db

RESULT 13
US-09-399-833A-245

Sequence 245, Application US/09999833A
Patent No. 6916648
GENERAL INFORMATION:
APPLICANT: Ashkenazi, Avi
APPLICANT: Baker Kevin P.
APPLICANT: Botstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Bacon, Dan
APPLICANT: Ferrara, Napoleon
APPLICANT: Filvaroff, Ellen
APPLICANT: Fong, Sherman
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerritsen, Mary B.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Hillan, Kenneth J.
APPLICANT: Kuo, Sophia S.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James J.
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Shelton, David L.
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P2630PIC65
CURRENT APPLICATION NUMBER: US/09/999,833A
CURRENT FILING DATE: 2001-10-24
PRIOR APPLICATION NUMBER: 09/918585
PRIOR FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/064249
PRIOR FILING DATE: 1997-11-03
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066364
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/077450
PRIOR FILING DATE: 1998-03-10
PRIOR APPLICATION NUMBER: 60/077632
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077641
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077649
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077791
PRIOR FILING DATE: 1998-03-12
PRIOR APPLICATION NUMBER: 60/078004
PRIOR FILING DATE: 1998-03-13
PRIOR APPLICATION NUMBER: 60/078886
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078936
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078939
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/079294
PRIOR FILING DATE: 1998-03-25
PRIOR APPLICATION NUMBER: 60/079656
PRIOR FILING DATE: 1998-03-26
PRIOR APPLICATION NUMBER: 60/079664
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079689
PRIOR FILING DATE: 1998-03-27

;; PRIOR APPLICATION NUMBER: 60/079663
;; PRIOR FILING DATE: 1998-03-27
;; PRIOR APPLICATION NUMBER: 60/079728
;; PRIOR FILING DATE: 1998-03-27
;; PRIOR APPLICATION NUMBER: 60/079786
;; PRIOR FILING DATE: 1998-03-27
;; PRIOR APPLICATION NUMBER: 60/079920
;; PRIOR FILING DATE: 1998-03-30
;; PRIOR APPLICATION NUMBER: 60/079923
;; PRIOR FILING DATE: 1998-03-30
;; PRIOR APPLICATION NUMBER: 60/080105
;; PRIOR FILING DATE: 1998-03-31
;; PRIOR APPLICATION NUMBER: 60/080107
;; PRIOR FILING DATE: 1998-03-31
;; PRIOR APPLICATION NUMBER: 60/080165
;; PRIOR FILING DATE: 1998-03-31
;; PRIOR APPLICATION NUMBER: 60/080194
;; PRIOR FILING DATE: 1998-03-31
;; PRIOR APPLICATION NUMBER: 60/080327
;; PRIOR FILING DATE: 1998-04-01
;; PRIOR APPLICATION NUMBER: 60/080328
;; PRIOR FILING DATE: 1998-04-01
;; PRIOR APPLICATION NUMBER: 60/080333
;; PRIOR FILING DATE: 1998-04-01
;; PRIOR APPLICATION NUMBER: 60/080334
;; PRIOR FILING DATE: 1998-04-01
;; PRIOR APPLICATION NUMBER: 60/081070
;; PRIOR FILING DATE: 1998-04-08
;; PRIOR APPLICATION NUMBER: 60/081049
;; PRIOR FILING DATE: 1998-04-08
;; PRIOR APPLICATION NUMBER: 60/081071
;; PRIOR FILING DATE: 1998-04-08
;; PRIOR APPLICATION NUMBER: 60/081195
;; PRIOR FILING DATE: 1998-04-08
;; PRIOR APPLICATION NUMBER: 60/081203
;; PRIOR FILING DATE: 1998-04-09
;; PRIOR APPLICATION NUMBER: 60/081229
;; PRIOR FILING DATE: 1998-04-09
;; PRIOR APPLICATION NUMBER: 60/081955
;; PRIOR FILING DATE: 1998-04-15
;; PRIOR APPLICATION NUMBER: 60/081817
;; PRIOR FILING DATE: 1998-04-15
;; PRIOR APPLICATION NUMBER: 60/081819
;; PRIOR FILING DATE: 1998-04-15
;; PRIOR APPLICATION NUMBER: 60/081952
;; PRIOR FILING DATE: 1998-04-15
;; PRIOR APPLICATION NUMBER: 60/081838
;; PRIOR FILING DATE: 1998-04-15
;; PRIOR APPLICATION NUMBER: 60/082568
;; PRIOR FILING DATE: 1998-04-21
;; PRIOR APPLICATION NUMBER: 60/082569
;; PRIOR FILING DATE: 1998-04-21
;; PRIOR APPLICATION NUMBER: 60/082704
;; PRIOR FILING DATE: 1998-04-22
;; PRIOR APPLICATION NUMBER: 60/082804
;; PRIOR FILING DATE: 1998-04-22
;; PRIOR APPLICATION NUMBER: 60/082700
;; PRIOR FILING DATE: 1998-04-22
;; PRIOR APPLICATION NUMBER: 60/082797
;; PRIOR FILING DATE: 1998-04-22
;; PRIOR APPLICATION NUMBER: 60/082796
;; PRIOR FILING DATE: 1998-04-23
;; PRIOR APPLICATION NUMBER: 60/083336
;; PRIOR FILING DATE: 1998-04-27
;; PRIOR APPLICATION NUMBER: 60/083332
;; PRIOR FILING DATE: 1998-04-28
;; PRIOR APPLICATION NUMBER: 60/083392
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083495
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083496
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083499

;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083545
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083554
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083558
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083559
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083500
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083742
;; PRIOR FILING DATE: 1998-04-30
;; PRIOR APPLICATION NUMBER: 60/084366
;; PRIOR FILING DATE: 1998-05-05
;; PRIOR APPLICATION NUMBER: 60/084414
;; PRIOR FILING DATE: 1998-05-06
;; PRIOR APPLICATION NUMBER: 60/084441
;; PRIOR FILING DATE: 1998-05-06
;; PRIOR APPLICATION NUMBER: 60/084637
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084639
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084640
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084598
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084600
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084627
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084643
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/085339
;; PRIOR FILING DATE: 1998-05-13
;; PRIOR APPLICATION NUMBER: 60/085338
;; PRIOR FILING DATE: 1998-05-13
;; PRIOR APPLICATION NUMBER: 60/085323
;; PRIOR FILING DATE: 1998-05-13
;; PRIOR APPLICATION NUMBER: 60/085582
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085700
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085689
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085579
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085580
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085573
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085704
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085697

Query Match 34.7%; Score 333.5; DB 2; Length 472;
Best Local Similarity 40.8%; Pred. No. 2.8e-28;
Matches 71; Conservative 24; Mismatches 60; Indels 19; Gaps 3;

QY 11 PLATATATLHTNRGDKILKLFGNHAKTYANFVGLAGKTKDYTSQASGSPSPFYDGAV 70
DB 9 PPTNGKVLTKTTAGDDIDBLMSKEAPKACNFIQLCLEA-----YYDNTT 53
QY 71 FRRVIOGFMIQGGDPGTGRGCP---GYKFADEFFHPELOFDKPYLLAMNAGSGTNGSQF 127
DB 54 FRRVVPGFIVQGGDPGTGTSGSGSESTYGAPPKDFEHLRLEFNRRGLVAMNAGSHDNGSQF 113
QY 128 FITVGTKPHLNRHRTTIFGEVYIDASQRYVEAISKRTADGNDREPTDVPVIESITI 181
DB 114 FFTLGRADELNNKHTIFGKVTGDTVYNNMLR-LSBVDIDDERPHNPHKIKSCV 166

RESULT 14

US-10-020-445A-245
Sequence 245, Application US/10020445A
Patent No. 6962797
GENERAL INFORMATION:
APPLICANT: Ashkenazi, Avi
APPLICANT: Baker Kevin P.
APPLICANT: Botstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Baton, Dan
APPLICANT: Ferrara, Napoleon
APPLICANT: Filvaroff, Ellen
APPLICANT: Fong, Sherman
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Ausetin L.
APPLICANT: Hillan, Kenneth J.
APPLICANT: Kijavlin, Ivar J.
APPLICANT: Kuo, Sophia S.
APPLICANT: Napier, Mary A.
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Shelton, David L.
APPLICANT: Stewart, Timothy A.
APPLICANT: Tuma, Daniel
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P2630P1C74
CURRENT APPLICATION NUMBER: US/10/020,445A
PRIOR FILING DATE: 2001-10-24
PRIOR APPLICATION NUMBER: 09/918585
PRIOR FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/064249
PRIOR FILING DATE: 1997-11-03
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066364
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/077450
PRIOR FILING DATE: 1998-03-10
PRIOR APPLICATION NUMBER: 60/077632
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077641
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077649
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077791
PRIOR FILING DATE: 1998-03-12
PRIOR APPLICATION NUMBER: 60/078004
PRIOR FILING DATE: 1998-03-13
PRIOR APPLICATION NUMBER: 60/078886
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078936
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078939
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/079294
PRIOR FILING DATE: 1998-03-25
PRIOR APPLICATION NUMBER: 60/079656
PRIOR FILING DATE: 1998-03-26
PRIOR APPLICATION NUMBER: 60/079664
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079689

PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079663
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079728
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079786
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079920
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/079923
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/080105
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080107
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080165
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080194
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080327
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080328
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080333
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080334
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/081070
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081049
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081071
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081195
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081203
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081229
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081955
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081817
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081819
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081952
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081838
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/082568
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082569
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082704
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082804
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082700
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082797
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082796
PRIOR FILING DATE: 1998-04-23
PRIOR APPLICATION NUMBER: 60/083336
PRIOR FILING DATE: 1998-04-27
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/083392
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083495
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083496
PRIOR FILING DATE: 1998-04-29

```
/ PRIOR APPLICATION NUMBER: 60/083499
/ PRIOR FILING DATE: 1998-04-29
/ PRIOR APPLICATION NUMBER: 60/083545
/ PRIOR FILING DATE: 1998-04-29
/ PRIOR APPLICATION NUMBER: 60/083554
/ PRIOR FILING DATE: 1998-04-29
/ PRIOR APPLICATION NUMBER: 60/083558
/ PRIOR FILING DATE: 1998-04-29
/ PRIOR APPLICATION NUMBER: 60/083559
/ PRIOR FILING DATE: 1998-04-29
/ PRIOR APPLICATION NUMBER: 60/083500
/ PRIOR FILING DATE: 1998-04-29
/ PRIOR APPLICATION NUMBER: 60/083742
/ PRIOR FILING DATE: 1998-04-30
/ PRIOR APPLICATION NUMBER: 60/084366
/ PRIOR FILING DATE: 1998-05-05
/ PRIOR APPLICATION NUMBER: 60/084414
/ PRIOR FILING DATE: 1998-05-06
/ PRIOR APPLICATION NUMBER: 60/084441
/ PRIOR FILING DATE: 1998-05-06
/ PRIOR APPLICATION NUMBER: 60/084637
/ PRIOR FILING DATE: 1998-05-07
/ PRIOR APPLICATION NUMBER: 60/084639
/ PRIOR FILING DATE: 1998-05-07
/ PRIOR APPLICATION NUMBER: 60/084640
/ PRIOR FILING DATE: 1998-05-07
/ PRIOR APPLICATION NUMBER: 60/084598
/ PRIOR FILING DATE: 1998-05-07
/ PRIOR APPLICATION NUMBER: 60/084600
/ PRIOR FILING DATE: 1998-05-07
/ PRIOR APPLICATION NUMBER: 60/084627
/ PRIOR FILING DATE: 1998-05-07
/ PRIOR APPLICATION NUMBER: 60/084643
/ PRIOR FILING DATE: 1998-05-07
/ PRIOR APPLICATION NUMBER: 60/085339
/ PRIOR FILING DATE: 1998-05-13
/ PRIOR APPLICATION NUMBER: 60/085338
/ PRIOR FILING DATE: 1998-05-13
/ PRIOR APPLICATION NUMBER: 60/085323
/ PRIOR FILING DATE: 1998-05-13
/ PRIOR APPLICATION NUMBER: 60/085582
/ PRIOR FILING DATE: 1998-05-15
/ PRIOR APPLICATION NUMBER: 60/085700
/ PRIOR FILING DATE: 1998-05-15
/ PRIOR APPLICATION NUMBER: 60/085689
/ PRIOR FILING DATE: 1998-05-15
/ PRIOR APPLICATION NUMBER: 60/085579
/ PRIOR FILING DATE: 1998-05-15
/ PRIOR APPLICATION NUMBER: 60/085580
/ PRIOR FILING DATE: 1998-05-15
/ PRIOR APPLICATION NUMBER: 60/085573
/ PRIOR FILING DATE: 1998-05-15
/ PRIOR APPLICATION NUMBER: 60/085704
/ PRIOR FILING DATE: 1998-05-15
/ PRIOR APPLICATION NUMBER: 60/085697
```

Query Match 34.7%; Score 333.5; DB 2; Length 472;

Best Local Similarity 40.8%; Pred. No. 2,8e-28; Matches 71; Conservative 24; Mismatches 60; Indels 19; Gaps 3;

```
QY 11 PLATATATHTNRGDIKIALFGNHAPKPTVANFVGLAQGTQDYSTONASGSPGPPYDGAVFHRIQGMIOG 70
DB 9 PPTNGKVLTKTGTAGDIDBLMSKAPKACRNFYOLCIEA-----YYDNTI 53
QY 71 FHRVIOGPMIOGDDPVTGRGCP---GYKPADBFHPELOPDKPYLLAMNAGPCTGSGOP 127
DB 54 FHRVVPGLTVQGGDPVTGSGGSGSIYGAAPKDFHRLFRNRGLVAMANAAGSHDNGSOP 113
QY 128 FITVGTPLNHRHTIFGEVIDAESORVVEAISKTATDGNDRPTDPVVISITI 181
DB 114 FTTLGADBLNKGHTIFGKVTGDTVYNNMLR-LSEVDIDDERPHNPHKIKSCV 166
```

```
RESULT 15
US-08-482-728A-8
/ Sequence 8, Application US/08482728A
/ Patent No. 5968802
/ GENERAL INFORMATION:
/ APPLICANT: Wang, Bruce
/ APPLICANT: Fisher, Joseph
/ APPLICANT: Payan, Donald
/ TITLE OF INVENTION: No. 5968802el Nuclear Cyclophilin
/ NUMBER OF SEQUENCES: 21
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Fleht, Hombach, Teet, Albritton
/ ADDRESSEE: & Herbert
/ STREET: Four Embarcadero Center, Suite 3400
/ CITY: San Francisco
/ STATE: California
/ COUNTRY: United States
/ ZIP: 94111-4187
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patent In Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/482,728A
/ FILING DATE: 07-JUN-1995
/ CLASSIFICATION: 435
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Silva, Robin M.
/ REGISTRATION NUMBER: 38,304
/ REFERENCE/DOCKET NUMBER: A-61230/DJB/RMS
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (415) 781-1989
/ TELEFAX: (415) 398-3249
/ TELEX: 910 277299
/ INFORMATION FOR SEQ ID NO: 8:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 123 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: unknown
/ TOPOLOGY: unknown
/ MOLECULE TYPE: protein
US-08-482-728A-8
```

Query Match 33.6%; Score 323; DB 1; Length 123;

Best Local Similarity 50.0%; Pred. No. 6,5e-28; Matches 70; Conservative 15; Mismatches 35; Indels 20; Gaps 4;

```
QY 24 GDIKIALFGNHAPKPTVANFVGLAQGTQDYSTONASGSPGPPYDGAVFHRIQGMIOG 83
DB 1 GDHHTLGFVEBCPTVENF-----CVHSRNG-----YNGHTFRIIKGFMIOG 45
QY 84 DPTGTGRGCP---GYKPADBFHPELOPDKPYLLAMNAGPCTGSGOPFITVGTPLNLR 140
DB 46 DPTGTGMBGSGSIWGGFEDBFHSTLHRDRPYLLSMNAGSNTNGSOPFITVGTPLNDR 105
QY 141 HTIFGEVIDAESORVVEAIS 160
DB 106 HTVFGRV--TKGMVYVQRIS 123
```

Search completed: April 14, 2006, 17:37:58
Job time: 17.829 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using SW model

Run on: April 14, 2006, 18:27:13 ; Search time 63.8683 Seconds
(without alignments)
1190.652 Million cell updates/sec

Title: US-10-620-246-12

Perfect score: 961
Sequence: 1 MADCDSVTMSPLATATATATLH.....ATDGNDRPTDPTVIESITIS 182

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_AA_Main:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pap:*
- 2: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pap:*
- 3: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pap:*
- 4: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pap:*
- 5: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pap:*
- 6: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	961	100.0	182	3	US-09-791-171-12 Sequence 12, Appl
2	961	100.0	182	3	US-09-804-980-12 Sequence 12, Appl
3	961	100.0	182	4	US-10-620-246-12 Sequence 12, Appl
4	716	74.5	187	3	US-09-738-626-3539 Sequence 3539, Ap
5	543.5	56.6	186	4	US-10-156-761-11863 Sequence 11863, A
6	407.5	42.4	466	4	US-10-474-776-694 Sequence 694, App
7	407.5	42.4	466	4	US-10-472-928-3150 Sequence 3150, Ap
8	403.5	42.0	472	5	US-10-617-320-4470 Sequence 4470, Ap
9	386	40.2	195	4	US-10-767-701-39256 Sequence 39256, A
10	386	40.2	651	4	US-10-437-963-144839 Sequence 144839,
11	384	40.0	160	4	US-10-424-599-168022 Sequence 168022,
12	384	39.9	160	4	US-10-437-963-196266 Sequence 196266,
13	381	39.6	166	4	US-10-767-701-61368 Sequence 61368, A
14	381	39.6	639	5	US-10-739-930-8097 Sequence 8097, Ap
15	380	39.5	616	4	US-10-424-599-212565 Sequence 212565,
16	376	39.1	235	4	US-10-425-115-237088 Sequence 237088,
17	375	39.0	246	4	US-10-425-115-286196 Sequence 286196,
18	370	38.5	165	4	US-10-425-115-237087 Sequence 237087,
19	359.5	37.4	420	4	US-10-158-057-327 Sequence 327, App
20	359.5	37.4	483	4	US-10-158-057-257 Sequence 257, App
21	359	37.4	176	6	US-11-097-143-33033 Sequence 33033, A
22	358.5	37.3	199	5	US-10-501-282-5676 Sequence 5676, Ap
23	358.5	37.3	202	5	US-10-501-282-5678 Sequence 5678, Ap
24	358.5	37.3	207	5	US-10-501-282-5680 Sequence 5680, Ap
25	352	36.6	164	4	US-10-425-115-278418 Sequence 278418,
26	352	36.6	166	4	US-10-028-072-8 Sequence 8, Appl
27	352	36.6	166	4	US-10-140-808-8 Sequence 8, Appl

28	352	36.6	166	4	US-10-121-049-8 Sequence 8, Appl
29	352	36.6	166	4	US-10-123-904-8 Sequence 8, Appl
30	352	36.6	166	4	US-10-140-470-8 Sequence 8, Appl
31	352	36.6	166	4	US-10-175-746-8 Sequence 8, Appl
32	352	36.6	166	4	US-10-176-918-8 Sequence 8, Appl
33	352	36.6	166	4	US-10-176-921-8 Sequence 8, Appl
34	352	36.6	166	4	US-10-137-865-8 Sequence 8, Appl
35	352	36.6	166	4	US-10-140-474-8 Sequence 8, Appl
36	352	36.6	166	4	US-10-142-431-8 Sequence 8, Appl
37	352	36.6	166	4	US-10-143-114-8 Sequence 8, Appl
38	352	36.6	166	4	US-10-142-419-8 Sequence 8, Appl
39	352	36.6	166	4	US-10-123-262-8 Sequence 8, Appl
40	352	36.6	166	4	US-10-142-423-8 Sequence 8, Appl
41	352	36.6	166	4	US-10-121-050-8 Sequence 8, Appl
42	352	36.6	166	4	US-10-141-755-8 Sequence 8, Appl
43	352	36.6	166	4	US-10-143-032-8 Sequence 8, Appl
44	352	36.6	166	4	US-10-123-108-8 Sequence 8, Appl
45	352	36.6	166	4	US-10-123-236-8 Sequence 8, Appl

ALIGNMENTS

```
RESULT 1
US-09-791-171-12
; Sequence 12, Application US/09791171
; Patent No. US2002009436A1
; GENERAL INFORMATION:
; APPLICANT: ANDERSEN, Peter
; APPLICANT: NIELSEN, Rikke
; APPLICANT: OSTTINGER, Thomas
; APPLICANT: RASMUSSEN, Peter Birk
; APPLICANT: ROSENKRANDS, Ida
; APPLICANT: WELDLINGH, Karin
; APPLICANT: FIORIO, Walter
; TITLE OF INVENTION: NUCLEIC ACIDS FRAGMENTS AND POLYPEPTIDE FRAGMENTS
; TITLE OF INVENTION: DERIVED FROM M. TUBERCULOSIS
; FILE REFERENCE: 670001-2002.1
; CURRENT APPLICATION NUMBER: US/09/791,171
; CURRENT FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 09/050,739
; PRIOR FILING DATE: 1998-03-30
; PRIOR APPLICATION NUMBER: 0376/97
; PRIOR FILING DATE: 1997-04-02
; PRIOR APPLICATION NUMBER: 1277/97
; PRIOR FILING DATE: 1997-11-10
; PRIOR APPLICATION NUMBER: 60/044,624
; PRIOR FILING DATE: 1997-04-18
; PRIOR APPLICATION NUMBER: 60/070,488
; PRIOR FILING DATE: 1998-01-05
; NUMBER OF SEQ ID NOS: 173
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12
; LENGTH: 182
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-09-791-171-12
Query Match 100.0%; Score 961; DB 3; Length 182;
Best Local Similarity 100.0%; Pred. No. 66-94;
Matches 182; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Cy 1 MADCDSVTMSPLATATATLHNRGDKIALFGHAPRTVANFGLAQGTQDYSTONASGG 60
Db 1 MADCDSVTMSPLATATATLHNRGDKIALFGHAPRTVANFGLAQGTQDYSTONASGG 60
Cy 61 PGGPFYDGAVFHRVIGFMIQGGDPTGTGCGGPGYFPADEFHDELQFDKXYLLAMNAGP 120
Db 61 PGGPFYDGAVFHRVIGFMIQGGDPTGTGCGGPGYFPADEFHDELQFDKXYLLAMNAGP 120
Cy 121 GTNGSOPFTTVGKTGPHLNRHRTTGGVITDAESQRVVEAISKTATDGNDRPTDPTVIESIT 180
Db 121 GTNGSOPFTTVGKTGPHLNRHRTTGGVITDAESQRVVEAISKTATDGNDRPTDPTVIESIT 180
```

QY 181 IS 182
Db 181 IS 182

RESULT 2

US-09-804-980-12
; Sequence 12, Application US/09804980
; Publication No. US20030147897A1
; GENERAL INFORMATION:
; APPLICANT: Statens Serum Institut
; APPLICANT: Andersen, Peter
; TITLE OF INVENTION: M. tuberculosis Antigens
; FILE REFERENCE: 670001-2002.4
; CURRENT APPLICATION NUMBER: US/09/804,980
; CURRENT FILING DATE: 2001-03-12
; NUMBER OF SEQ ID NOS: 257
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 12
; LENGTH: 182
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-09-804-980-12

Query Match 100.0%; Score 961; DB 3; Length 182;
Best Local Similarity 100.0%; Pred. No. 6e-94;
Matches 182; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MADCDVTSPLATATATLTHTNRGDIKIALFGNHAKPTVANFVGLAQGTQDYSTONASGG 60
Db 1 MADCDVTSPLATATATLTHTNRGDIKIALFGNHAKPTVANFVGLAQGTQDYSTONASGG 60
QY 61 PSQPFYDGAVFHRVIOGFMIQSGDPTGTGRGPGYKFADEFFHPELOFDKPYLLAMNAGP 120
Db 61 PSQPFYDGAVFHRVIOGFMIQSGDPTGTGRGPGYKFADEFFHPELOFDKPYLLAMNAGP 120
QY 121 GTNGSQFFITVGKTPHLNRHHTIFGEVIDAESQRVVEAISKTATDGNDRPTDPVIESIT 180
Db 121 GTNGSQFFITVGKTPHLNRHHTIFGEVIDAESQRVVEAISKTATDGNDRPTDPVIESIT 180

QY 181 IS 182
Db 181 IS 182

RESULT 3

US-10-620-246-12
; Sequence 12, Application US/10620246
; Publication No. US20040115211A1
; GENERAL INFORMATION:
; APPLICANT: Andersen, Peter
; APPLICANT: NIELSEN, Rikke
; APPLICANT: OETTINGER, Thomas
; APPLICANT: RASMUSSEN, Peter Birk
; APPLICANT: ROSENKRANDS, Ida
; APPLICANT: WELDJING, Karin
; APPLICANT: FLORIO, Walter
; TITLE OF INVENTION: NUCLEIC ACIDS FRAGMENTS AND POLYPEPTIDE FRAGMENTS
; TITLE OF INVENTION: DERIVED FROM M. TUBERCULOSIS
; FILE REFERENCE: 670001-2002.1A
; CURRENT APPLICATION NUMBER: US/10/620,246
; CURRENT FILING DATE: 2003-07-15
; PRIOR APPLICATION NUMBER: 09/050,739
; PRIOR FILING DATE: 1998-03-30
; PRIOR APPLICATION NUMBER: 0376/97
; PRIOR FILING DATE: 1997-04-02
; PRIOR APPLICATION NUMBER: 1277/97
; PRIOR FILING DATE: 1997-11-10
; PRIOR APPLICATION NUMBER: 60/044,624
; PRIOR FILING DATE: 1997-04-18
; PRIOR APPLICATION NUMBER: 60/070,488
; PRIOR FILING DATE: 1998-01-05

; PRIOR APPLICATION NUMBER: 10/138,473
; PRIOR FILING DATE: 2002-05-02
; PRIOR APPLICATION NUMBER: 09/791,171
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 09/415,884
; PRIOR FILING DATE: 1999-10-08
; PRIOR APPLICATION NUMBER: 60/116,673
; PRIOR FILING DATE: 1999-01-21
; PRIOR APPLICATION NUMBER: 1281/98
; PRIOR FILING DATE: 1998-10-08
; NUMBER OF SEQ ID NOS: 173
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12
; LENGTH: 182
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-10-620-246-12

Query Match 100.0%; Score 961; DB 4; Length 182;
Best Local Similarity 100.0%; Pred. No. 6e-94;
Matches 182; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MADCDVTSPLATATATLTHTNRGDIKIALFGNHAKPTVANFVGLAQGTQDYSTONASGG 60
Db 1 MADCDVTSPLATATATLTHTNRGDIKIALFGNHAKPTVANFVGLAQGTQDYSTONASGG 60
QY 61 PSQPFYDGAVFHRVIOGFMIQSGDPTGTGRGPGYKFADEFFHPELOFDKPYLLAMNAGP 120
Db 61 PSQPFYDGAVFHRVIOGFMIQSGDPTGTGRGPGYKFADEFFHPELOFDKPYLLAMNAGP 120
QY 121 GTNGSQFFITVGKTPHLNRHHTIFGEVIDAESQRVVEAISKTATDGNDRPTDPVIESIT 180
Db 121 GTNGSQFFITVGKTPHLNRHHTIFGEVIDAESQRVVEAISKTATDGNDRPTDPVIESIT 180
QY 181 IS 182
Db 181 IS 182

RESULT 4

US-09-738-626-3539
; Sequence 3539, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 3539
; LENGTH: 187
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-3539

Query Match 74.5%; Score 716; DB 3; Length 187;

Best Local Similarity 79.9%; Pred. No. 8e-68;
Matches 135; Conservative 11; Mismatches 23; Indels 0; Gaps 0;

QY 14 TATATLTNRGDIKIALFGNHA PKTVANFVGLA QGTDKYSTONASGGSPGPFYGA VHR 73
DB 18 TATATLTNRGDIKIALFGNHA PKTVANFVGLA QGTDKYSTONASGGSPGPFYGA VHR 77
QY 74 VIOGFMIQGGDPGTGGRGPGYKAFDEPHPELOFDPKPYLLAMANA GPTNGSQPFITV GK 133
DB 78 VIOGFMIQGGDPGTGGRGPGYKAFDEPHPELOFDPKPYLLAMANA GPTNGSQPFITV TP 137
QY 134 TPHLNRRHTT FGEVYDAESQRYVAISKTATDGNDRPTDPVVISITIT 182
DB 138 TPHLNRRHTT FGEVYDAESQRYVAISKTATDGNDRPTDPVVISITIT 186

RESULT 5

US-10-156-761-11863
Sequence 11863, Application US/10156761
Publication No. US20030119018A1

GENERAL INFORMATION:
APPLICANT: OMURA, SATOSHI
APPLICANT: IKEDA, HARUO
APPLICANT: ISHIKAWA, TUN
APPLICANT: HORIKAWA, HIROSHI
APPLICANT: SHIBA, TADAYOSHI
APPLICANT: SAKAKI, YOSHIYUKI
APPLICANT: HATTORI, MASAHIRA
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-262
CURRENT APPLICATION NUMBER: US/10/156,761
PRIOR FILING DATE: 2002-05-29
PRIOR APPLICATION NUMBER: JP 2001-204089
PRIOR FILING DATE: 2001-05-30
PRIOR APPLICATION NUMBER: JP 2001-272697
PRIOR FILING DATE: 2001-08-02
NUMBER OF SEQ ID NOS: 15109
SEQ ID NO 11863
LENGTH: 186
TYPE: PRT
ORGANISM: Streptococcus avermitilis
US-10-156-761-11863

Query Match 56.6%; Score 543.5; DB 4; Length 186;
Best Local Similarity 65.3%; Pred. No. 1.9e-49;
Matches 109; Conservative 17; Mismatches 36; Indels 3; Gaps 3;

QY 17 ATLHTNRGDIKIALFGNHA PKTVANFVGLA QGTDKYSTONASGGSPGPFYGA VHRVI 75
DB 7 ATLHTNRGDIKIALFGNHA PKTVANFVGLA QGTDKYSTONASGGSPGPFYGA VHRVI 65
QY 76 QGFMIOGGDPGTGGRGPGYKAFDEPHPELOFDPKPYLLAMANA GPTNGSQPFITV GKTP 135
DB 66 SGMIOGGDPGTGGRGPGYKAFDEPHPELOFDPKPYLLAMANA GPTNGSQPFITV APTA 135
QY 136 HLNRHTT FGEVYDAESQRYVAISKTATDGNDRPTDPVVISITIT 181
DB 126 HLNRHTT FGEVYDAESQRYVAISKTATDGNDRPTDPVVISITIT 172

RESULT 6

US-10-474-776-694
Sequence 694, Application US/10474776
Publication No. US20040110181A1

GENERAL INFORMATION:
APPLICANT: Wyeth
TITLE OF INVENTION: NOVEL STREPTOCOCCUS PNEUMONIAE OPEN READING FRAMES ENCODING POLYH
FILE REFERENCE: AM100649-PCF
CURRENT APPLICATION NUMBER: US/10/474,776
CURRENT FILING DATE: 2003-10-14
NUMBER OF SEQ ID NOS: 752
SOFTWARE: PatentIn version 3.1

SEQ ID NO 694
LENGTH: 466
TYPE: PRT
ORGANISM: Streptococcus pneumoniae
US-10-474-776-694

Query Match 42.4%; Score 407.5; DB 4; Length 466;
Best Local Similarity 45.0%; Pred. No. 2e-34;
Matches 94; Conservative 23; Mismatches 41; Indels 51; Gaps 7;

QY 3 DCSVNSPLATATLTNRGDIKIALFGNHA PKTVANFVGLA QGTDKYSTONASGGSP 62
DB 277 DITV-BGPL-ATIKNHGDLRIKLPFHAPKTVANFVGLA QGTDKYSTONASGGSP 317
QY 63 GPFYDGA VHRVIQGFMIQGGDPGTGGRGPGYKAFDEPHPELOFDPKPYLLAMANA G 119
DB 318 -GYDDGVIFHRITIKDPMIOGGDPGTGGRGPGYKAFDEPHPELOFDPKPYLLAMANA G 375
QY 120 PGTNGSQPFI-----TVKTPHLNRRHTT FGEVYDAES 152
DB 376 PNTNGSQPFIYONQHLPYSKKEITRGWMPPIAIVANOGGTPLHRRHTVFGQLADEAS 435
QY 153 QRYVAISKTATDGNDRPTDPVVISITIT 181
DB 436 YAVLDAIAAVETGAMDKPVEDVITETI 464

RESULT 7

US-10-472-928-3150
Sequence 3150, Application US/10472928
Publication No. US2005020813A1

GENERAL INFORMATION:
APPLICANT: CHIRON SPA
APPLICANT: THE INSTITUTE FOR GENOMIC RESEARCH
TITLE OF INVENTION: STREPTOCOCCUS PNEUMONIAE PROTEINS AND NUCLEIC ACIDS
FILE REFERENCE: P026926WO
CURRENT APPLICATION NUMBER: US/10/472,928
PRIOR FILING DATE: 2003-09-26
PRIOR APPLICATION NUMBER: GB-0107658.7
PRIOR FILING DATE: 2001-03-27
NUMBER OF SEQ ID NOS: 4979
SOFTWARE: SeqWIn9, version 1.03
SEQ ID NO 3150
LENGTH: 466
TYPE: PRT
ORGANISM: Streptococcus pneumoniae
FEATURE:
OTHER INFORMATION: cof family protein/peptidyl-prolyl cis-trans isomerase, cyclophi
OTHER INFORMATION: Cellular location: cytoplasm
OTHER INFORMATION: Similar to strain R6 sequence 15903436 (O.R+01)
US-10-472-928-3150

Query Match 42.4%; Score 407.5; DB 5; Length 466;
Best Local Similarity 45.0%; Pred. No. 2e-34;
Matches 94; Conservative 23; Mismatches 41; Indels 51; Gaps 7;

QY 3 DCSVNSPLATATLTNRGDIKIALFGNHA PKTVANFVGLA QGTDKYSTONASGGSP 62
DB 277 DITV-BGPL-ATIKNHGDLRIKLPFHAPKTVANFVGLA QGTDKYSTONASGGSP 317
QY 63 GPFYDGA VHRVIQGFMIQGGDPGTGGRGPGYKAFDEPHPELOFDPKPYLLAMANA G 119
DB 318 -GYDDGVIFHRITIKDPMIOGGDPGTGGRGPGYKAFDEPHPELOFDPKPYLLAMANA G 375
QY 120 PGTNGSQPFI-----TVKTPHLNRRHTT FGEVYDAES 152
DB 376 PNTNGSQPFIYONQHLPYSKKEITRGWMPPIAIVANOGGTPLHRRHTVFGQLADEAS 435
QY 153 QRYVAISKTATDGNDRPTDPVVISITIT 181
DB 436 YAVLDAIAAVETGAMDKPVEDVITETI 464

```

RESULT 8
US-10-617-320-4470
/ Sequence 4470, Application US/10617320
/ Publication No. US20050136404AI
/ GENERAL INFORMATION:
/ APPLICANT: Lylyn A Doucette-Stamm and David Bush
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID
/ SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE
/ THERAPEUTICS
/ NUMBER OF SEQUENCES: 5206
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: GENOME THERAPEUTICS CORPORATION
/ STREET: 100 Beaver Street
/ CITY: Waltham
/ STATE: Massachusetts
/ COUNTRY: USA
/ ZIP: 02354
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: CD-ROM ISO9660
/ COMPUTER: <Unknown>
/ OPERATING SYSTEM: <Unknown>
/ SOFTWARE: <Unknown>
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/10/617,320
/ FILING DATE: 10-Jul-2003
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US/09/107,433
/ FILING DATE: 30-Jun-1998
/ APPLICATION NUMBER: 60/ 085131
/ FILING DATE: May 12, 1998
/ APPLICATION NUMBER: 60/051553
/ FILING DATE: July 2, 1997
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Arinello, Pamela Deneke
/ REGISTRATION NUMBER: 40, 489
/ REFERENCE/DOCKET NUMBER: GTC-011
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (781)893-5007
/ TELEFAX: (781)893-8277
/ INFORMATION FOR SEQ ID NO: 4470:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 472 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ HYPOTHEICAL: YES
/ ORIGINAL SOURCE:
/ ORGANISM: Streptococcus pneumoniae
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (B) LOCATION 1..472
/ SEQUENCE DESCRIPTION: SEQ ID NO: 4470:
US-10-617-320-4470

Query Match 42.0%; Score 403.5; DB 5; Length 472;
Beac Local Similarity 44.5%; Pred. No. 5.4e-34;
Matches 93; Conservative 23; Mismatches 42; Indels 51; Gaps 7;

3 DCDSTNSPLATATATLTITNRGDIKIALFGNHAPPTVAFVGLAQGTQVSTQNASGGS 62
Db 283 DIERV-EEPL-----ATITKNGDRIKILFPEHAKRTYANFVSL-----SKD----- 323
63 GPFYDGAVFHNVIGFMIGGSDPTGTGRGP--GYKPADEFHDELQFDKPYLLIAMNAG 119
Db 324 -GYVDGVIHFHIIIDFMIGGDDPTGTGIGESIVGESPEDEFSBEL-YNIGALSMANAG 361
QY 120 PGTNGSGPFI-----TVGKTPHIANRRTIFGVIAAS 152
Db 382 PNTGSGPFIYQONHLPYSKKEITRGWPEPIABIIYANQGCTPHIDRRHTVFGQADAEAS 441
QY 153 QRVEAISKATDGDNDPTDPVIESITI 181
Db 442 YAVLDVIAAVETGANDKPEVDVIEITETI 470

```

```

RESULT 9
US-10-767-701-39256
Sequence 39256, Application US/10767701
Publication No. US2004017284A1
GENERAL INFORMATION:
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(5353)B
CURRENT FILING DATE: 2004-01-29
NUMBER OF SEQ ID NOS: 63128
SEQ ID NO 39256
LENGTH: 195
TYPE: PRT
ORGANISM: Sorghum bicolor
FEATURE:
OTHER INFORMATION: Clone ID: SORBI_28MAV03-C10248_1.pep
US-10-767-701-39256

Query Match      40.2%; Score 386; DB 4; Length 195;
Beet Local Similarity .46.3%; Pred.No.1,2e-32;
Matches      82; Conservative    23; Mismatches     46; Gaps       5;

14 TATAT-----LHTNRGDIKIALFGNHA PKTVANFVGILAQGT KDYSTONASGGSPGPYD   67
33 TATSLPDLNLVHSTMSGDIIHLRLYPEBCKTVENF-----TTCRNG----YYD   77

Qy      68 GAVERHYVQGSMIGQGDPTGCRGRGP---GYCFADFEHLELPDKRYLLAMNAGPGTG   124
Db      78 NLIFHRVYKGFMIDGTDPGLSGDTGGQSITWTEFEDBFHSLRHDRFTLSMANAGPGTTG   137

Oy      125 SQPFITWGKTPHLNRRHTIFGEVIDAESQRVEAISKATPDGNDREPTDPEVIESITI   181
Db      138 SQPFITTVATWLDNKHTVFGRV--KGMADVQQLKTKTDINDKRYGVKLIINIV   192

RESULT 10
US-10-437-963-144839
Sequence 144839, Application US/10437963
Publication No. US20040123343A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
APPLICANT: Wu, Wei
APPLICANT: Bouharov, Andrey A.
APPLICANT: Barbazuk, Brad
APPLICANT: Li, Ping
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53221)B
CURRENT FILING DATE: 2003-05-14
NUMBER OF SEQ ID NOS: 204966
SEQ ID NO 144839
LENGTH: 651
TYPE: PRT
ORGANISM: Oryza sativa
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT4530_45616C.1.pep
US-10-437-963-144839

Query Match      40.2%; Score 386; DB 4; Length 651;
Beet Local Similarity .46.3%; Pred.No.6.1e-32;
Matches      82; Conservative    23; Mismatches     46; Indels     26; Gaps       5;

14 TATATL-----HTNRGDIKIALFGNHA PKTVANFVGILAQGT KDYSTONASGGSPGPYD   67

```

```

Db      489 TATSTLEPDMVMTSMWDHRLRLYPEBCPKTVEF-----TTHGRNG---YTD 533
Qy      68 GAVFHRVIOGFMIQGGDPYTGGRGP---GKYFADBFHPELQFDKPYLLAMANAGPGTNG 124
      534 NLIFHRVYKGFMIQYGPDLGDTGGOSIWMGRFEDFHKSLRHDPPTLSMANAGPGTNG 593
Qy      125 SQPFITVGTPTPLNRHRTIFGEVIDAESQRYVEAISKTATDGNDRPTDPVIESITI 181
      594 SQPFITVATPMTDNKRTVGRVY--KGMVYVQOIEKVKTDKDKPFYQDVXILNVTV 648

```

RESULT 11

```

US-10-424-599-168022
; Sequence 168022, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 168022
; LENGTH: 160
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_12273C.1.pep
US-10-424-599-168022

```

```

Query Match      40.0%; Score 384; DB 4; Length 160;
Best Local Similarity 47.9%; Pred. No. 1.Se-32;
Matches 81; Conservative 22; Mismatches 46; Indels 20; Gaps 4;

```

```

Qy      16 TATLHNRGDIKIALFGNHAPKTVANFVGLAQGTQKYSTQNASGSPGSPYDGAVFHRVI 75
      2 SVTLHNLGDIKCEVFCDOQPRTAENFALC-----ASG-----YTDGTFHRNI 46
Qy      76 QGFMIQGGDPYTGGRGP---GKYFADBFHPELQFDKPYLLAMANAGPGTNGSQFTTVG 132
      47 KGFMIQGGDPYTGGRGKSTISWKKFADBFHPELQFDKPYLLAMANAGPGTNGSQFTTVG 106
Db      133 KTFHLNRHRTIFGEVIDAESQRYVEAISKTATDGNDRPTDPVIESITI 181
      107 KQPHLNGHYTVFAKVI--HGFEVLDLMEKAKQTGDRPLAELRLNRVTI 153

```

RESULT 12

```

US-10-437-963-196266
; Sequence 196266, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barabak, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 196266
; LENGTH: 160

```

```

; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_92133C.1.pep
US-10-437-963-196266

```

```

Query Match      39.9%; Score 383; DB 4; Length 160;
Best Local Similarity 47.3%; Pred. No. 2e-32;
Matches 80; Conservative 24; Mismatches 45; Indels 20; Gaps 4;

```

```

Qy      16 TATLHNRGDIKIALFGNHAPKTVANFVGLAQGTQKYSTQNASGSPGSPYDGAVFHRVI 75
      2 SVTLHNLGDIKCEVFCDOQPRTAENFALC-----ASG-----YTDGTFHRNI 46
Qy      76 QGFMIQGGDPYTGGRGP---GKYFADBFHPELQFDKPYLLAMANAGPGTNGSQFTTVG 132
      47 KGFMIQGGDPYTGGRGKSTISWKKFADBFHPELQFDKPYLLAMANAGPGTNGSQFTTVG 106
Db      133 KTFHLNRHRTIFGEVIDAESQRYVEAISKTATDGNDRPTDPVIESITI 181
      107 KQPHLNGHYTVFAKVI--HGFEVLDLMEKAKQTGDRPLAELRLNRVTI 153

```

RESULT 13

```

US-10-767-701-61368
; Sequence 61368, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; NUMBER OF SEQ ID NOS: 2004-01-29
; SEQ ID NO 61368
; LENGTH: 166
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: 9304935.pep
US-10-767-701-61368

```

```

Query Match      39.6%; Score 381; DB 4; Length 166;
Best Local Similarity 47.3%; Pred. No. 3.4e-32;
Matches 80; Conservative 23; Mismatches 46; Indels 20; Gaps 4;

```

```

Qy      16 TATLHNRGDIKIALFGNHAPKTVANFVGLAQGTQKYSTQNASGSPGSPYDGAVFHRVI 75
      8 SVTLHNLGDIKCEVFCDOQPRTAENFALC-----ASG-----YTDGTFHRNI 52
Qy      76 QGFMIQGGDPYTGGRGP---GKYFADBFHPELQFDKPYLLAMANAGPGTNGSQFTTVG 132
      53 KGFMIQGGDPYTGGRGKSTISWKKFADBFHPELQFDKPYLLAMANAGPGTNGSQFTTVG 112
Db      133 KTFHLNRHRTIFGEVIDAESQRYVEAISKTATDGNDRPTDPVIESITI 181
      113 KQPHLNGHYTVFAKVI--HGFEVLDLMEKAKQTGDRPLAELRLNRVTI 159

```

RESULT 14

```

US-10-739-930-8097
; Sequence 8097, Application US/10739930
; Publication No. US20040216190A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH
; TITLE OF INVENTION: PLANTS AND USES THEREOF FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-21(53377)B
; CURRENT APPLICATION NUMBER: US/10/739,930
; NUMBER OF SEQ ID NOS: 2003-12-18

```

```
/ NUMBER OF SEQ ID NOS: 11088
/ SEQ ID NO 8097
/ LENGTH: 639
/ TYPE: PRT
/ ORGANISM: Zea mays
/ FEATURE:
/ NAME/KEY: unsure
/ LOCATION: (1)..(639)
/ OTHER INFORMATION: unsure at all Xaa locations
/ FEATURE:
/ OTHER INFORMATION: Clone ID: ZEAMN-23APR03-C38338_1.p
US-10-739-930-8097
```

```
Query Match          39.6%; Score 381; DB 5; Length 639;
Best Local Similarity 45.8%; Pred. No. 2e-31;
Matches 81; Conservative 23; Mismatches 47; Indels 26; Gaps 5;
```

```
QY 14 TATAT-----LTNRGDIKIALFGNHAFTVANFVGLAOGTKDYSTONASGSPSGPYD 67
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 477 TATSLPDNLVLTSMGDIHLRLYPEECPTVENF-----TTHCRXG---YVD 521

QY 68 GAVFHVIOGFMIQGGDPGTGGRGP--GYKFADEFHPELOFDPKPYLLAMNAGPTNG 124
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 522 NLIFHRYIKGFMIQTGDDPLGDTGGQSIWGTREDEFHKSLRHDPPTLSMANAGPTNG 581

QY 125 SQPFITVGKTPHLNRHTTIFGEVYDAESQRYVEAISKTAIDGNDRPTDPVIESITI 181
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 582 SQPFITTVATFPMIDNKHIVFGRVY--KGMDEVQIEKVTDKNDKPYQDVKILNVTV 636
```

RESULT 15

```
US-10-424-599-212565
/ Sequence 212565, Application US/10424599
/ Publication No. US20040031072A1
/ GENERAL INFORMATION:
/ APPLICANT: La Rosa Thomas J
/ APPLICANT: Kovalic David K
/ APPLICANT: Zhou Yinhua
/ APPLICANT: Cao Yongwei
/ TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
/ TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
/ FILE REFERENCE: 38-21(53223)B
/ CURRENT APPLICATION NUMBER: US/10/424,599
/ FILING DATE: 2003-04-28
/ NUMBER OF SEQ ID NOS: 285684
/ SEQ ID NO 212565
/ LENGTH: 616
/ TYPE: PRT
/ ORGANISM: Glycine max
/ FEATURE:
/ OTHER INFORMATION: Clone ID: PAT_MRT3847_33973C.1.pep
US-10-424-599-212565
```

```
Query Match          39.5%; Score 380; DB 4; Length 616;
Best Local Similarity 46.4%; Pred. No. 2.5e-31;
Matches 83; Conservative 21; Mismatches 53; Indels 22; Gaps 5;
```

```
QY 6 SVTNSPLATATATLTNRGDIKIALFGNHAFTVANFVGLAOGTKDYSTONASGSPSGPYD 65
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 454 SVTTS--LPDNVLTHTTMDIHKLYPEECPTVENF-----TTHCRNG-----Y 496

QY 66 YDGAVFHRYIOGFMIQGGDPGTGGRGP--GYKFADEFHPELOFDPKPYLLAMNAGPT 122
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 497 YDMLIFHRYIKGFMIQTGDDPLGDTGGQSIWGTREDEFHKSLRHDPPTLSMANAGPT 556

QY 123 NGSQPFITVGKTPHLNRHTTIFGEVYDAESQRYVEAISKTAIDGNDRPTDPVIESITI 181
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 557 NGSQPFITTVATFPMIDNKHIVFGRV--AKGMDVVOAIEKVTDKNDKPYQDVKILNVTV 613
```

Search completed: April 14, 2006, 18:40:38
Job time : 64.9683 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: April 14, 2006, 18:37:38 ; Search time 7.79883 Seconds
(without alignments)
991.565 Million cell updates/sec

Title: US-10-620-246-12

Perfect score: 961
Sequence: 1 MADCDSVTNSPLATATATLH.....ATDGNRPDPVIESITIS 182

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 217505 seqs, 42489236 residues

Total number of hits satisfying chosen parameters: 217505

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA New:
1: /SIDS5/ptodata/2/pubpaa/US06_NEM_PUB.pep:*
2: /SIDS5/ptodata/2/pubpaa/US06_NEM_PUB.pep:*
3: /SIDS5/ptodata/2/pubpaa/US07_NEM_PUB.pep:*
4: /SIDS5/ptodata/2/pubpaa/PCT_NEM_PUB.pep:*
5: /SIDS5/ptodata/2/pubpaa/US09_NEM_PUB.pep:*
6: /SIDS5/ptodata/2/pubpaa/US10_NEM_PUB.pep:*
7: /SIDS5/ptodata/2/pubpaa/US11_NEM_PUB.pep:*
8: /SIDS5/ptodata/2/pubpaa/US60_NEM_PUB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	352	36.6	166	US-10-131-826A-8	Sequence 8, Appl1
2	352	36.6	166	US-10-517-151-6	Sequence 6, Appl1
3	352	36.6	166	US-10-973-115B-8	Sequence 8, Appl1
4	352	36.6	166	US-11-290-153-8	Sequence 8, Appl1
5	348	36.2	164	US-11-024-959-310	Sequence 310, App
6	347	36.1	164	US-11-096-568A-21480	Sequence 21480, A
7	345	35.9	163	US-11-024-959-435	Sequence 435, App
8	333.5	34.7	472	US-10-195-883-98	Sequence 98, App1
9	333.5	34.7	472	US-10-195-883-98	Sequence 98, App1
10	333.5	34.7	472	US-10-195-883-98	Sequence 98, App1
11	333.5	34.7	472	US-10-195-883-98	Sequence 98, App1
12	332	34.5	164	US-11-096-568A-33153	Sequence 33153, A
13	332	34.5	164	US-11-024-959-442	Sequence 442, App
14	330	34.3	140	US-11-096-568A-21481	Sequence 21481, A
15	331	33.4	140	US-11-096-568A-21481	Sequence 21481, A
16	318	33.1	147	US-11-096-568A-33154	Sequence 33154, A
17	315	32.8	143	US-11-096-568A-6908	Sequence 6908, App
18	312	32.5	140	US-11-096-568A-33155	Sequence 33155, A
19	312	32.5	502	US-11-024-959-301	Sequence 301, App
20	308	32.0	504	US-11-096-568A-29206	Sequence 29206, A
21	297.5	31.0	361	US-11-024-959-438	Sequence 438, App
22	292.5	30.4	252	US-11-096-568A-1494	Sequence 1494, App
23	292.5	30.4	257	US-11-096-568A-1493	Sequence 1493, App
24	292.5	30.4	264	US-11-096-568A-1492	Sequence 1492, App
25	291.5	30.3	248	US-11-096-568A-20638	Sequence 20638, A

26	291.5	30.3	262	7	US-11-024-959-302	Sequence 302, App
27	291	30.0	208	7	US-11-227-238-44	Sequence 44, App1
28	288	30.0	257	7	US-11-024-959-303	Sequence 303, App
29	287.5	29.9	262	7	US-11-024-959-430	Sequence 430, App
30	283.5	29.5	361	7	US-11-024-959-313	Sequence 313, App
31	282.5	29.4	172	7	US-11-024-959-305	Sequence 305, App
32	282.5	29.4	172	7	US-11-024-959-319	Sequence 319, App
33	282.5	29.4	172	7	US-11-024-959-323	Sequence 323, App
34	281	29.2	663	7	US-11-024-959-314	Sequence 314, App
35	278.5	28.5	204	7	US-11-024-959-433	Sequence 433, App
36	276.5	28.8	172	6	US-10-991-285-879	Sequence 879, App
37	274.5	28.6	203	7	US-11-024-959-308	Sequence 308, App
38	274.5	28.6	213	7	US-11-096-568A-21250	Sequence 21250, A
39	274.5	28.6	239	7	US-11-096-568A-21249	Sequence 21249, A
40	274.5	28.6	242	7	US-11-096-568A-21248	Sequence 21248, A
41	273.5	28.5	172	7	US-11-024-959-306	Sequence 306, App
42	273.5	28.5	172	7	US-11-024-959-425	Sequence 425, App
43	272.5	28.4	172	7	US-11-024-959-317	Sequence 317, App
44	272.5	28.4	172	7	US-11-024-959-424	Sequence 424, App
45	272.5	28.4	172	7	US-11-024-959-426	Sequence 426, App

ALIGNMENTS

RESULT 1
US-10-131-826A-8
; Sequence 8, Application US/10131826A
; Publication No. US20050245730A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerlitsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C128
; CURRENT APPLICATION NUMBER: US/10/131,826A
; CURRENT FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: 60/049911
; PRIOR FILING DATE: 1997-06-18
; PRIOR APPLICATION NUMBER: 60/056974
; PRIOR FILING DATE: 1997-08-26
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059115
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059117
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059122
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059184
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059352
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/059588
; PRIOR FILING DATE: 1997-09-19
; Remaining Prior Application data removed - See File Wrapper or PALM.

[illegible]

```

; CURRENT APPLICATION NUMBER: US/11/024,959
; CURRENT FILING DATE: 2004-12-30
; PRIOR APPLICATION NUMBER: 60/533,036
; PRIOR FILING DATE: 2003-12-30
; NUMBER OF SEQ ID NOS: 782
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 310
; LENGTH: 164
; TYPE: PRT
; ORGANISM: Eucalyptus sp.
US-11-024-959-310

Query Match          36.2%; Score 348; DB 7; Length 164;
Beet Local Similarity 43.8%; Pred. No. 6.2e-29;
Matches 71; Conservative 22; Mismatches 49; Indels 20; Gaps 3

18 TLATNRGDIKIALPGHNAFKTVANFVGLAQCTKDYSTONASGSGSPFYDGAVFERRVIOG 77
|||:::||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
13 TLSTSMGSPFYELVFKAPFRTSRNFIELSR-----RGYYDNVKKFRIIKD 57

QY 78 FMIDGGDPTGTGRGSP---GYKPADBFHPELOFDKPYLLAMNANGPGTNGSOFFTTVGKT 134
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 58 FIVGGSPPTGGRGSGESYIGKKFDEDEIKPELKHGAGCILSMNANGPNTNGSOFFITLAPC 117
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

QY 135 PHLNRBHTEGEVVIDASQGVVAELSKATDGGNDRPDPVVI 176
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 118 PSLDGKHTIFGRV--CRGMEILIKLQAGSVQTDNNRPIHDVCI 157
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

```

```

RESULT 6
US-11-096-568A-21480
; Sequence 21480, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
APPLICANT: Alexandrov, Nikolai et al.
TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
FILE REFERENCE: 2750-1592PUS2
CURRENT APPLICATION NUMBER: US/11/096,568A
CURRENT FILING DATE: 2005-04-01
NUMBER OF SEQ ID NOS: 34471
SEQ ID NO 21480
LENGTH: 164
TYPE: PRT
ORGANISM: Zea mays subsp. mays
FEATURES:
NAME/KEY: misc feature
LOCATION: (1)..(164)
OTHER INFORMATION: Ceres Seq. ID no. 12404718
US-11-096-568A-21480

Query Match          36.1%; Score 347; DB 7; Length 164;
Best Local Similarity 44.0%; Pred. No. 7, 9e-28;
Matches 73; Conservative 21; Mismatches 52; Indels 20; Gaps 3;

Cy      14 TATATTHTNRGDIKIALFGNHAPTTVANFVGLAQGTQDYSTONASGSGSPFYDGAVFHR 73
       | | | | : | | | : | | | | | | | | | | | | | | | | | | | | | | |
Db      9 TPBATLTETSMGAVSVEMYYRHAPTCRRNFVELAR-----RCGYDNVIFHR 53

Cy      74 VIOGFMILOGGDPPTGTRGGR---GYKPADSEPHPELOFDKPYLLAMANAGPGTNGSQFFIT 130
       ||| ||||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     54 IIKPIYVGGDPTGTGGBSIYGAKTFDELIKTKITKTAGACILSNANAAGPPTNGSQFFIT 113

Cy      131 VGKTPHLNRRHTTIFGEVIDAESQRVEAISKTATDGANDRPTDPVVI 176
       | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     114 LAPCQSIDGKHITFGRV--CRGMEIVRLSGSVQTDKNDRPITHEVKI 157

RESULT 7
US-11-024-959-435
; Sequence 435, Application US/11024959
; Publication No. US20060010516A1
; GENERAL INFORMATION:
```

```
/ APPLICANT: FORSTER, RICHARD L.
/ APPLICANT: CONNETT, MARIE B.
/ APPLICANT: EMERSON, SARAH JANE
/ APPLICANT: GRIGOR, MURRAY ROBERT
/ APPLICANT: HIGGINS, COLLEEN M.
/ APPLICANT: LUND, STEVEN TROY
/ APPLICANT: MAGUSIN, ANDREAS
/ APPLICANT: KODRZYCKI, BOB
/ TITLE OF INVENTION: CELL CYCLE GENES AND RELATED METHODS
/ FILE REFERENCE: 04463-0360
/ CURRENT APPLICATION NUMBER: US/11/024,959
/ PRIOR FILING DATE: 2004-12-30
/ PRIOR APPLICATION NUMBER: 60/533,036
/ PRIOR FILING DATE: 2003-12-30
/ NUMBER OF SEQ ID NOS: 782
/ SOFTWARE: Patentin version 3.3
/ SEQ ID NO 435
/ LENGTH: 163
/ TYPE: PRT
/ ORGANISM: Pinus radiata
US-11-024-959-435

Query Match          35.9%; Score 345; DB 7; Length 163;
Best Local Similarity 43.9%; Pred. No. 1,2e-27; Mismatches 53; Indels 20; Gaps 3;
Matches 75; Conservative 23;

QY 14 TATATATGTRGDIKIALFGNHAPKTYANFVGLAOGTKDYSTQNASGGSPGYDGAVFHR 73
DB 8 TREVLETSMGPTVELYHNAKPKTCRNFELSR-----RGYNNKVFHR 52
QY 74 VIQGFMIQGGDPGTGGRGP---GYKPADEFHPELOPKDYLLAMANAAGGTNGSQFFIT 130
DB 53 VIKDFVVOGGDPGTGGRGSGESYIGAPRFEDEITRDLKHTGAGILSMANAGFNTNGSQFFIS 112
QY 131 VGRTPHLMNRHTIFGEVIDAESQRYVEAISKTAIDGNDPRTDVPVIESITI 181
DB 113 LAPTPLWLDKHTIFGRV--CKGMDVVRGLGNVOTDKNDRIHDVKILRTTV 161

RESULT 8
US-10-194-487-98
/ Sequence 98, Application US/10194487
/ Publication No. US20060074226A1
/ GENERAL INFORMATION:
/ APPLICANT: Baker, Kevin P.
/ APPLICANT: Chen, Jian
/ APPLICANT: Desnoyers, Luc
/ APPLICANT: Goddard, Audrey
/ APPLICANT: Godowski, Paul J.
/ APPLICANT: Gurney, Austin L.
/ APPLICANT: Pan, James
/ APPLICANT: Smith, Victoria
/ APPLICANT: Matanabe, Colin K.
/ APPLICANT: Wood, William I.
/ APPLICANT: Zhang, Zemin
/ TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
/ TITLE OF INVENTION: ACIDS ENCODING THE SAME
/ FILE REFERENCE: P3430R1C312
/ CURRENT APPLICATION NUMBER: US/10/194,487
/ CURRENT FILING DATE: 2002-07-12
/ PRIOR APPLICATION NUMBER: 10/052586
/ PRIOR FILING DATE: 2002-01-15
/ PRIOR APPLICATION NUMBER: 60/059263
/ PRIOR FILING DATE: 1997-09-18
/ PRIOR APPLICATION NUMBER: 60/059266
/ PRIOR FILING DATE: 1997-09-18
/ PRIOR APPLICATION NUMBER: 60/062250
/ PRIOR FILING DATE: 1997-10-17
/ PRIOR APPLICATION NUMBER: 60/063120
/ PRIOR FILING DATE: 1997-10-24
/ PRIOR APPLICATION NUMBER: 60/063121
/ PRIOR FILING DATE: 1997-10-24
/ PRIOR APPLICATION NUMBER: 60/063486
```

```
/ PRIOR FILING DATE: 1997-10-21
/ PRIOR APPLICATION NUMBER: 60/063540
/ PRIOR FILING DATE: 1997-10-28
/ PRIOR APPLICATION NUMBER: 60/063541
/ PRIOR FILING DATE: 1997-10-28
/ PRIOR APPLICATION NUMBER: 60/063544
/ PRIOR FILING DATE: 1997-10-28
/ Prior Application data removed - See file wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 612
/ SEQ ID NO 98
/ LENGTH: 472
/ TYPE: PRT
/ ORGANISM: Homo Sapien
US-10-195-883-98

Query Match          34.7%; Score 333.5; DB 6; Length 472;
Best Local Similarity 40.8%; Pred. No. 7.4e-26;
Matches 71; Conservative 24; Mismatches 60; Indels 19; Gaps 3;

QY 11 PLATATATLHTNRGDIKIALFGNHAPKTYANFVGLAOGTKDYSTQNASGGSPGYDGA 70
DB 9 PPTNGRVLLKTTAGDIDIELMSKAPKACRNFQLCLEA-----YYDNTI 53
QY 71 FHRVIOGFMIQGGDPGTGGRGP---GYKPADEFHPELOPKDYLLAMANAAGGTNGSQF 127
DB 54 FHRVVGFTVVOGGDPGTGSGGESYIGAPFKDBFHRSLFRNRGLVAMANAAGSHDNGSQF 113
QY 128 FIVTGPHLMNRHTIFGEVIDAESQRYVEAISKTAIDGNDPRTDVPVIESITI 181
DB 114 FFTLGRADLNMNRHTIFGKYTDVTVNMLR--LSVDIDDERHNRHKIKSCV 166

RESULT 9
US-10-195-883-98
/ Sequence 98, Application US/10195883
/ Publication No. US20060073544A1
/ GENERAL INFORMATION:
/ APPLICANT: Baker, Kevin P.
/ APPLICANT: Chen, Jian
/ APPLICANT: Desnoyers, Luc
/ APPLICANT: Goddard, Audrey
/ APPLICANT: Godowski, Paul J.
/ APPLICANT: Gurney, Austin L.
/ APPLICANT: Pan, James
/ APPLICANT: Smith, Victoria
/ APPLICANT: Matanabe, Colin K.
/ APPLICANT: Wood, William I.
/ APPLICANT: Zhang, Zemin
/ TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
/ TITLE OF INVENTION: ACIDS ENCODING THE SAME
/ FILE REFERENCE: P3430R1C323
/ CURRENT APPLICATION NUMBER: US/10/195,883
/ CURRENT FILING DATE: 2002-07-15
/ Prior Application removed - See file wrapper or Palm
/ NUMBER OF SEQ ID NOS: 612
/ SEQ ID NO 98
/ LENGTH: 472
/ TYPE: PRT
/ ORGANISM: Homo Sapien
US-10-195-883-98

Query Match          34.7%; Score 333.5; DB 6; Length 472;
Best Local Similarity 40.8%; Pred. No. 7.4e-26;
Matches 71; Conservative 24; Mismatches 60; Indels 19; Gaps 3;

QY 11 PLATATATLHTNRGDIKIALFGNHAPKTYANFVGLAOGTKDYSTQNASGGSPGYDGA 70
DB 9 PPTNGRVLLKTTAGDIDIELMSKAPKACRNFQLCLEA-----YYDNTI 53
QY 71 FHRVIOGFMIQGGDPGTGGRGP---GYKPADEFHPELOPKDYLLAMANAAGGTNGSQF 127
DB 54 FHRVVGFTVVOGGDPGTGSGGESYIGAPFKDBFHRSLFRNRGLVAMANAAGSHDNGSQF 113
```

QY 128 FIVGKTPHNRRTTIGEVDAESQVVEAISKATATGNDPRPTDPAVVISITTI 181
 DB 114 FFTLGRADDELNNKHTIFGKVTGDTVYNNMLR-LSEVDIDDDRRPHNPHKIXSCEV 166

RESULT 10

US-10-195-888-98
 / Sequence 98, Application US/10195888
 / Publication No. US20060072545A1
 / GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.
 APPLICANT: Chen, Jilan
 APPLICANT: Desnoyers, Luc
 APPLICANT: Goddard, Audrey J.
 APPLICANT: Godowski, Paul J.
 APPLICANT: Gurney, Austin L.
 APPLICANT: Pan, James
 APPLICANT: Smith, Victoria
 APPLICANT: Watanabe, Collin K.
 APPLICANT: Wood, William I.
 APPLICANT: Zhang, Zhenli
 TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 TITLE OF INVENTION: ACIDS ENCODING THE SAME
 FILE REFERENCE: P3430R1C324
 CURRENT APPLICATION NUMBER: US/10/195,888
 PRIOR FILING DATE: 2002-07-15
 NUMBER OF SEQ ID NOS: 612
 SEQ ID NO 98
 LENGTH: 472
 TYPE: PRT
 ORGANISM: Homo Sapien
 US-10-195-888-98

Query Match 34.7%; Score 333.5; DB 6; Length 472;
 Best Local Similarity 40.8%; Pred. No. 7.4e-26;
 Matches 71; Conservative 24; Mismatches 60; Indels 19; Gaps 3;

QY 11 PLATATATLTHTNGDIALFGNHPKTVANFVGLAGTODYSTQNASGGSPGYDGA 70
 DB 9 PPTNGKVLTKTTAGDIDIELMSKAPACRNFIOLCLEA-----YDNTI 53
 QY 71 FHRVIOGFMIQGGPPTGSGRGP---GYKPADBFHPELOFDKPYLLAMANAGPPTNGSQF 127
 DB 54 FHRVVPFIVQGGPPTGSGGSIYCAPFDEBHSRLRFRGLVAMANAGSHDNGSQF 113
 QY 128 FIVGKTPHNRRTTIGEVDAESQVVEAISKATATGNDPRPTDPAVVISITTI 181
 DB 114 FFTLGRADDELNNKHTIFGKVTGDTVYNNMLR-LSEVDIDDDRRPHNPHKIXSCEV 166

RESULT 11

US-10-195-889-98
 / Sequence 98, Application US/10195889
 / Publication No. US20060074227A1
 / GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.
 APPLICANT: Chen, Jilan
 APPLICANT: Desnoyers, Luc
 APPLICANT: Goddard, Audrey J.
 APPLICANT: Godowski, Paul J.
 APPLICANT: Gurney, Austin L.
 APPLICANT: Pan, James
 APPLICANT: Smith, Victoria
 APPLICANT: Watanabe, Collin K.
 APPLICANT: Wood, William I.
 APPLICANT: Zhang, Zhenli
 TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 TITLE OF INVENTION: ACIDS ENCODING THE SAME
 FILE REFERENCE: P3430R1C329
 CURRENT APPLICATION NUMBER: US/10/195,889
 PRIOR FILING DATE: 2002-07-15
 NUMBER OF SEQ ID NOS: 612
 SEQ ID NO 98
 LENGTH: 472
 TYPE: PRT
 ORGANISM: Homo Sapien
 US-10-195-889-98

QY 128 FIVGKTPHNRRTTIGEVDAESQVVEAISKATATGNDPRPTDPAVVISITTI 181
 DB 114 FFTLGRADDELNNKHTIFGKVTGDTVYNNMLR-LSEVDIDDDRRPHNPHKIXSCEV 166

Query Match 34.7%; Score 333.5; DB 6; Length 472;
 Best Local Similarity 40.8%; Pred. No. 7.4e-26;
 Matches 71; Conservative 24; Mismatches 60; Indels 19; Gaps 3;

QY 11 PLATATATLTHTNGDIALFGNHPKTVANFVGLAGTODYSTQNASGGSPGYDGA 70
 DB 9 PPTNGKVLTKTTAGDIDIELMSKAPACRNFIOLCLEA-----YDNTI 53
 QY 71 FHRVIOGFMIQGGPPTGSGRGP---GYKPADBFHPELOFDKPYLLAMANAGPPTNGSQF 127
 DB 54 FHRVVPFIVQGGPPTGSGGSIYCAPFDEBHSRLRFRGLVAMANAGSHDNGSQF 113
 QY 128 FIVGKTPHNRRTTIGEVDAESQVVEAISKATATGNDPRPTDPAVVISITTI 181
 DB 114 FFTLGRADDELNNKHTIFGKVTGDTVYNNMLR-LSEVDIDDDRRPHNPHKIXSCEV 166

RESULT 12

US-11-096-568A-33153
 / Sequence 33153, Application US/11096568A
 / Publication No. US20060048240A1
 / GENERAL INFORMATION:

APPLICANT: Alexandrov, Nikolai et al.
 TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
 FILE REFERENCE: 2750-1592PUS2
 CURRENT APPLICATION NUMBER: US/11/096,568A
 CURRENT FILING DATE: 2005-04-01
 NUMBER OF SEQ ID NOS: 34471
 SEQ ID NO 33153
 LENGTH: 164
 TYPE: PRT
 ORGANISM: Arabidopsis thaliana
 FEATURES:
 NAME/KEY: misc_feature
 LOCATION: (1)..(164)
 OTHER INFORMATION: Ceres Seq. ID no. 13602270
 US-11-096-568A-33153

Query Match 34.5%; Score 332; DB 7; Length 164;
 Best Local Similarity 41.4%; Pred. No. 2.6e-26;
 Matches 67; Conservative 26; Mismatches 49; Indels 20; Gaps 3;

QY 18 TLHTNGDIALFGNHPKTVANFVGLAGTODYSTQNASGGSPGYDGA 77
 DB 13 TLTSTWGPFTVEWYKHSPTCRNFLELSR-----RGYDNLPHRIVKO 57
 QY 78 FMIQGGPPTGSGRGP---GYKPADBFHPELOFDKPYLLAMANAGPPTNGSQF 134
 DB 58 FIVQGGPPTGSGGSIYCAPFDEBHSRLRFRGLVAMANAGPPTNGSQF 117
 QY 135 PHNRRTTIGEVDAESQVVEAISKATATGNDPRPTDPAVVISITTI 176
 DB 118 PSIDGKHTIGRV--CRGMEVIRKLSGVQDNDTRPIHEVXI 157

RESULT 13

US-11-024-959-442
 / Sequence 442, Application US/11024959
 / Publication No. US20060010516A1
 / GENERAL INFORMATION:

APPLICANT: FORSTER, RICHARD L.
 APPLICANT: CONNETT, MARIE B.
 APPLICANT: EMERSON, SARAH JANE
 APPLICANT: GRIGOR, MURRAY ROBERT
 APPLICANT: HIGGINS, COLLEEN M.

```
/ APPLICANT: LUND, STEVEN TROY
/ APPLICANT: MAGUSIN, ANDREAS
/ APPLICANT: KODRZYCKI, BOB
/ TITLE OF INVENTION: CELL CYCLE GENES AND RELATED METHODS
/ FILE REFERENCE: 04463-0360
/ CURRENT APPLICATION NUMBER: US/11/024,959
/ PRIOR FILING DATE: 2004-12-30
/ PRIOR APPLICATION NUMBER: 60/533,036
/ NUMBER OF SEQ ID NOS: 782
/ SOFTWARE: PatentIn version 3.3
/ SEQ ID NO 442
/ LENGTH: 373
/ TYPE: PRT
/ ORGANISM: Pinus radiata
US-11-024-959-442
```

```
Query Match      34.5%; Score 332; DB 7; Length 373;
Best Local Similarity 48.4%; Pred. No. 7.7e-26;
Matches 78; Conservative 18; Mismatches 49; Indels 16; Gaps 6;
```

```
QY 24 GDIKIALFGNHAKPTVANFVGLAOGTKDYSTONASGGSPGFYDGAVFHRYIOGPMIQ 81
DB 23 GRIVFSLYADLVKTKAENFALCTGKGI-----GSKGKPLCYAGSGFHRVIKGFMCQ 75
QY 82 GGDPT-GTGRGSP---GYKPADBFHPELOPDKPYLLAMANAQPGTNGSQPFITVGTGTPHL 137
DB 76 GGFPTANGTGGISYIGKEFDEAF-VKHTKPFLLSMANAGKOTNGSQPFITVGTGTPHL 134
QY 138 NRRHTIFGEVIDAESQRYVEAISKTATDGNDRPTDPVVIS 178
DB 135 DDGHVVFGEVIKGRS--IVRAIENYPTASGDVPTSPITISA 173
```

RESULT 14

```
US-11-096-568A-21481
/ Sequence 21481, Application US/11096568A
/ Publication No. US20060048240A1
/ GENERAL INFORMATION:
/ APPLICANT: Alexandrov, Nikolai et al.
/ TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
/ FILE REFERENCE: 2750-1592PUS2
/ CURRENT APPLICATION NUMBER: US/11/096,568A
/ CURRENT FILING DATE: 2005-04-01
/ NUMBER OF SEQ ID NOS: 34471
/ SEQ ID NO 21481
/ LENGTH: 147
/ TYPE: PRT
/ ORGANISM: Zea mays subsp. mays
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (1)-(147)
/ OTHER INFORMATION: Cereals Seq. ID no. 12404719
US-11-096-568A-21481
```

```
Query Match      34.3%; Score 330; DB 7; Length 147;
Best Local Similarity 44.2%; Pred. No. 3.5e-26;
Matches 69; Conservative 20; Mismatches 47; Indels 20; Gaps 3;
```

```
QY 24 GDIKIALFGNHAKPTVANFVGLAOGTKDYSTONASGGSPGFYDGAVFHRYIOGPMIQ 83
DB 2  GAIVSEMYRYHAKPTCNFVELAR-----RGYYDNVIFRRIIKDFIVOGG 46
QY 84 DPTGTGRGSP---GYKPADBFHPELOPDKPYLLAMANAQPGTNGSQPFITVGTGTPHLNR 140
DB 47 DPTGTGRGSGSIYGAKFEDIKTELKRTGAGILSMANAGPNTNGSQPFITLAPCQSLDGK 106
QY 141 HTIFGEVIDAESQRYVEAISKTATDGNDRPTDPVVI 176
DB 107 HTIFGRV--CRGMEIVKRIKLSVQTDKNDRIHEVKI 140
```

```
RESULT 15
US-11-096-568A-21482
/ Sequence 21482, Application US/11096568A
/ Publication No. US20060048240A1
/ GENERAL INFORMATION:
/ APPLICANT: Alexandrov, Nikolai et al.
/ TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
/ FILE REFERENCE: 2750-1592PUS2
/ CURRENT APPLICATION NUMBER: US/11/096,568A
/ CURRENT FILING DATE: 2005-04-01
/ NUMBER OF SEQ ID NOS: 34471
/ SEQ ID NO 21482
/ LENGTH: 140
/ TYPE: PRT
/ ORGANISM: Zea mays subsp. mays
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (1)-(140)
/ OTHER INFORMATION: Cereals Seq. ID no. 12404720
US-11-096-568A-21482
```

```
Query Match      33.4%; Score 321; DB 7; Length 140;
Best Local Similarity 45.3%; Pred. No. 2.7e-25;
Matches 68; Conservative 18; Mismatches 44; Indels 20; Gaps 3;
```

```
QY 30 LFGNHAKPTVANFVGLAOGTKDYSTONASGGSPGFYDGAVFHRYIOGPMIOGGDPTGTG 89
DB 1  MYRYHAKPTCNFVELAR-----RGYYDNVIFRRIIKDFIVOGGDPGTGTG 45
QY 90 RGGP---GYKPADBFHPELOPDKPYLLAMANAQPGTNGSQPFITVGTGTPHLNRHTITGE 146
DB 46 RGGESIYGAKFEDIKTELKRTGAGILSMANAGPNTNGSQPFITLAPCQSLDGKHTIFGR 105
QY 147 VIDAESQRYVEAISKTATDGNDRPTDPVVI 176
DB 106 V--CRGMEIVKRIKLSVQTDKNDRIHEVKI 133
```

```
Search completed: April 14, 2006, 18:42:15
Job time : 7.79883 secs
```

GenCore version 5.1.7
Copyright (c) 1993 - 2006 BioCeleration Ltd.

OM protein - protein search, using sw model

Run on: April 14, 2006, 17:03:40 ; Search time 102.677 Seconds
(without alignments)
1133.996 Million cell updates/sec

Title: US-10-620-246-16

Perfect score: 1348
Sequence: 1 MNNLYRDLAPVTEAAWAAR.....LQSTLTFLCYTAASVALSH 265

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_21.*
1: geneseqp1980s.*
2: geneseqp1990s.*
3: geneseqp2000s.*
4: geneseqp2001s.*
5: geneseqp2002s.*
6: geneseqp2003as.*
7: geneseqp2003bs.*
8: geneseqp2004s.*
9: geneseqp2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1348	100.0	265	AAW72891	Mycobacte
2	1348	100.0	265	AAW721908	AAW721908
3	147	10.9	287	ABM95900	Abm95900 M. xanthu
4	108	8.0	1574	AAE37002	AAE37002 Micromono
5	103	7.6	8805	ABBB67112	ABBB67112 Drosophi
6	100	7.4	19	AAW72898	AAW72898 Mycobacte
7	100	7.4	19	AAW721915	AAW721915 N-termi
8	99.5	7.4	755	ABBB91978	ABBB91978 Herbidida
9	96	7.1	555	ADN26519	ADN26519 Bacteri
10	96	7.1	579	ADN26762	ADN26762 Bacteri
11	96	7.1	614	ADN28471	ADN28471 Bacteri
12	96	7.1	621	ABBB66212	ABBB66212 Drosophi
13	95	7.0	4038	ABG99871	ABG99871 S. cinna
14	94.5	7.0	853	AAAB26917	AAAB26917 large sub
15	94.5	7.0	853	ABU62618	ABU62618 Pseudomon
16	93.5	6.9	263	ABO71472	ABO71472 Pseudomon
17	92.5	6.9	583	ADV49538	ADV49538 OIL-assoc
18	92	6.8	728	ADX92106	ADX92106 Plant ful
19	91.5	6.8	2194	AAW40114	AAW40114 Human pol
20	91	6.8	381	ABO61838	ABO61838 Klebsiell
21	91	6.8	383	AAAG29504	AAAG29504 Arabidops
22	91	6.8	441	ADM26564	ADM26564 Hyperther
23	91	6.8	442	ABU52687	ABU52687 Human sig
24	91	6.8	469	AAAG29503	AAAG29503 Arabidops

25	91	6.8	469	ADN72195	ADN72195 Thale cre
26	91	6.8	472	ADT59335	ADT59335 Plant pol
27	91	6.8	702	ABU24332	ABU24332 Protein e
28	90.5	6.7	245	ADR24082	ADR24082 Pyrococcu
29	90.5	6.7	345	ADR24079	ADR24079 Pyrococcu
30	90.5	6.7	526	ADS29018	ADS29018 Bacteri
31	90.5	6.7	1220	ABG24318	ABG24318 Novel hum
32	90	6.7	328	ABU43515	ABU43515 Protein e
33	90	6.7	464	ADS26263	ADS26263 Bacteri
34	90	6.7	505	ADS26229	ADS26229 Bacteri
35	90	6.7	505	ADS27020	ADS27020 Bacteri
36	90	6.7	530	ABO71541	ABO71541 Pseudomon
37	90	6.7	792	ABU36373	ABU36373 Pseudomon
38	90	6.7	792	ABU38650	ABU38650 Protein e
39	90	6.7	792	ABU18816	ABU18816 Pseudomon
40	89.5	6.6	268	ABO84045	ABO84045 Pseudomon
41	89.5	6.6	416	ADN22084	ADN22084 Bacteri
42	89.5	6.6	416	ADN24843	ADN24843 Bacteri
43	89.5	6.6	530	ABP66019	ABP66019 Bifidobac
44	89.5	6.6	539	AAU97744	AAU97744 Mycobacte
45	89.5	6.6	539	AAU98656	AAU98656 Mycobacte

ALIGNMENTS

RESULT 1
AAW72891 standard; protein; 265 AA.
ID AAW72891
AC AAW72891;
AD 21-JAN-1999 (first entry)
DE Mycobacterium tuberculosis antigen CFP29.
DX Mycobacterium tuberculosis; antigen; vaccine; immunological; immunogen;
KW Mycobacterium tuberculosis; antigen; vaccine; immunological; immunogen;
XX Infection.
XX Mycobacterium tuberculosis.
OS Mycobacterium tuberculosis.
XX W09844119-A1.
PN 08-OCT-1998.
PD 01-APR-1998; 98WO-DK000132.
PF 02-APR-1997; 97DK-00000376.
PR 18-APR-1997; 97US-0044624P.
PR 10-NOV-1997; 97DK-00001277.
PR 05-JAN-1998; 98US-0070488P.
XX (STAT-) STATENS SERUM INST.
XX Andersen P, Nielsen R, Rosenkrands I, Weidling K, Rasmussen PB;
PI Oettinger T, Florio W;
XX WPI; 1998-542705/46.
DR N-PSDB; AAV63922.
XX New isolated mycobacteria polypeptides and nucleic acids - used for
PT developing products for the diagnosis of or vaccination against
PT mycobacterial infections, particularly tuberculosis.
PS Claim 1, Page 140; 163pp; English.
XX The present sequence represents a Mycobacterium tuberculosis protein.
CC Products from the present invention, which describes protein fragments
CC and nucleic acid fragments derived from M. tuberculosis, can be used in
CC the detection of and prevention of mycobacterial infections. In
CC particular, the proteins and nucleic acids can be used for the diagnosis
CC of or vaccination against tuberculosis caused by M. tuberculosis, M.
CC africanum or M. bovis

XX SQ Sequence 265 AA;
 Query Match 100.0%; Score 1348; DB 2; Length 265;
 Best Local Similarity 100.0%; Pred. No. 2.3e-130;
 Matches 265; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNLYRDLAPVTEAAWAEIEEAARTFKRHIAGRVVDSDPGPYTAAVSTGRLLDVKA 60
 DB 1 MNLYRDLAPVTEAAWAEIEEAARTFKRHIAGRVVDSDPGPYTAAVSTGRLLDVKA 60
 QY 61 PTNGVIAHLRASKEPLVRLRVPFTLSRNEIDVVERGSKDSWEPVKEAAKGLAFVEERTIF 120
 DB 61 PTNGVIAHLRASKEPLVRLRVPFTLSRNEIDVVERGSKDSWEPVKEAAKGLAFVEERTIF 120
 QY 121 EGYSAASIEGIRASASNPALTLPEDPREIPDVISQALSELRLAGVGPYSVLLSADVYTK 180
 DB 121 EGYSAASIEGIRASASNPALTLPEDPREIPDVISQALSELRLAGVGPYSVLLSADVYTK 180
 QY 181 VSESTDHGYPIREHNLRLVDGDIIMAPALDGAFLVLTTRGGDFDLQGTDVVAIGYASHDT 240
 DB 181 VSESTDHGYPIREHNLRLVDGDIIMAPALDGAFLVLTTRGGDFDLQGTDVVAIGYASHDT 240
 QY 241 TERLYIQEITLTFCTYAEASVALSH 265
 DB 241 TERLYIQEITLTFCTYAEASVALSH 265

RESULT 2
 ID AAY21908 standard; protein; 265 AA.
 XX AAY21908;
 AC 06-SEP-1999 (first entry)
 DT Amino acid sequence of antigen CFP29.
 DE Immunogenic; Mycobacterium tuberculosis; immune response; infection;
 KW tuberculous; fusion polypeptide; T-cell epitope; BSAT-6; MPT59; TB;
 KW pharmaceutical; vaccination; M. africanum; M. bovis; CFP2A; CFP30A;
 KW CFP2B; CFP19; CFP27; CFP30A; RD1-ORF; CFP10A; CFP16; CFP19; CFP23;
 KW CFP25A; CFP30B; CFP7B.
 XX Mycobacterium tuberculosis.
 OS
 PN MO9924577-AI.
 XX 20-MAY-1999.
 PD 08-OCT-1998; 98MO-DK000438.
 PF 10-NOV-1997; 97DK-00001277.
 PR 05-JUN-1998; 98US-0070488P.
 PR 01-APR-1998; 98MO-DK000132.
 XX (STAR-) STATENS SERUM INST.
 PA Andersen P, Skjot R;
 PI WPI; 1999-347282/29.
 DR N-PSDB; AAX81008.
 XX New immunogenic fragment of Mycobacterium tuberculosis.
 PT Example 3; Page 60-61; 265pp; English.
 PS The invention describes a substantially pure immunogenic polypeptide
 CC fragment (1) from Mycobacterium tuberculosis that is able to evoke a
 CC protective immune response against infections by mycobacteria belonging
 CC to the tuberculosis complex. The invention provides a (1) fusion
 CC polypeptide comprising at least one polypeptide fragment (1) and at least
 CC one fusion partner; (2) a fusion polypeptide fragment comprising a T-cell

CC epitope from M. tuberculosis protein BSAT-6, or MPT59 and a second
 CC different amino acid sequence from M. tuberculosis, and/or including a
 CC sequence which protects the first amino acid sequence from in vivo
 CC degradation or post-translational processing; (3) a nucleic acid fragment
 CC that encodes the above polypeptides. The polypeptides and nucleic acid
 CC are useful as pharmaceuticals, for diagnosis of and as antigens for
 CC vaccination against TB caused by Mycobacterium tuberculosis, africanum or
 CC bovis. The polypeptides are also useful for diagnosing ongoing or
 CC previous sensitization in an animal with bacteria belonging to the
 CC tuberculosis complex. The invention also describes the use of CFP7A or
 CC CFP30A or a T-cell epitope of for the induction of a strong immune
 CC response in a mammal; use of CFP7B, CFP19 or MPT59-BSAT6 or a T-cell
 CC epitope of for diagnosis of TB in a mammal by performing a DRH type skin
 CC test; use of CFP27, CFP30A, RD1-ORF2, RD1-ORF3, RD1-ORF5, MPT59-BSAT6,
 CC BSAT6-MPT59, CFP10A, CFP16, CFP19, CFP23, CFP25A, CFP30B, CFP7B or a T-
 CC cell epitope of for the preparation of an immunological composition; and
 CC for the preparation of a subunit vaccine

XX SQ Sequence 265 AA;
 Query Match 100.0%; Score 1348; DB 2; Length 265;
 Best Local Similarity 100.0%; Pred. No. 2.3e-130;
 Matches 265; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNLYRDLAPVTEAAWAEIEEAARTFKRHIAGRVVDSDPGPYTAAVSTGRLLDVKA 60
 DB 1 MNLYRDLAPVTEAAWAEIEEAARTFKRHIAGRVVDSDPGPYTAAVSTGRLLDVKA 60
 QY 61 PTNGVIAHLRASKEPLVRLRVPFTLSRNEIDVVERGSKDSWEPVKEAAKGLAFVEERTIF 120
 DB 61 PTNGVIAHLRASKEPLVRLRVPFTLSRNEIDVVERGSKDSWEPVKEAAKGLAFVEERTIF 120
 QY 121 EGYSAASIEGIRASASNPALTLPEDPREIPDVISQALSELRLAGVGPYSVLLSADVYTK 180
 DB 121 EGYSAASIEGIRASASNPALTLPEDPREIPDVISQALSELRLAGVGPYSVLLSADVYTK 180
 QY 181 VSESTDHGYPIREHNLRLVDGDIIMAPALDGAFLVLTTRGGDFDLQGTDVVAIGYASHDT 240
 DB 181 VSESTDHGYPIREHNLRLVDGDIIMAPALDGAFLVLTTRGGDFDLQGTDVVAIGYASHDT 240
 QY 241 TERLYIQEITLTFCTYAEASVALSH 265
 DB 241 TERLYIQEITLTFCTYAEASVALSH 265

RESULT 3
 ID ABM95900 standard; protein; 287 AA.
 XX ABM95900;
 AC 02-JUN-2005 (first entry)
 DT M. xanthus protein sequence, seq id 15099.
 DE Transgenic plant; DNA replication; gene regulation; gene expression.
 KW Myxococcus xanthus.
 OS
 PN US6833447-BI.
 XX 21-DEC-2004.
 PD 10-JUL-2001; 2001US-00902540.
 PF 10-JUL-2000; 2000US-0217883P.
 PR (MONS) MONSANTO TECHNOLOGY LLC.
 XX Goldman BS, Hinkle GJ, Slater SC, Wiegand RC;
 PI WPI; 2005-028716/03.

Query	Best Local Similarity	8.0%; Score 108; DB 6; Length 1574;
Matches	86; Conservative 30; Mismatches 100; Indels 94; Gaps 20;	
Qy	6 RDL---APVTEAAMAEIETEAARTFKKH-----IAGR-RVVDVSD--PGGPTAAVSTGR 54	
Db	506 RDLIAAPPEIDLDASAGIALTAASVDPDRNAVVTAAAGRPQALDLDLAGSEPPRAYTGG- 564	
Qy	55 LIDVKAFTN-----GVIAHLRASKPLV---RLRVPTLSRN---EIDDER 94	
Db	565 ---VAAPTGTFTVFPFGQTHMGAGADLLDQSPVFAESMRRCQALSAITDWKLGVR 621	
Qy	95 GSKDSD-----WEPTKEA-----KKLAFVEDRTTFEGYSAASIEGYSASNP 139	
Db	622 GAAGSPEDLDVDVLIQFVMAVMVSLAQWMSLGLVEPDAAV--CHSQGE--IAAAVWCGA 676	
Qy	140 LTLPEPDRELPDVISOALSELRLAGDGPYSVTLASADVTKVSETDHYPIREHNLRLV 199	
Db	677 LTLPDARVVA--LRSQVIGRV--LSGGGMAASVQL-----PAEVAAGRLA 718	
Qy	200 DGDIIWPAIDGAFV-----LITRGGDFDLQ-----GTDV-----ALGYASHDTPTER 243	
Db	719 ----AAAGRLDVAANVGPGSTVSGAADAATVELVEAFPAEDVAVRRLPVDYASHSTQVDR 774	
Qy	244 LYIQETLITFL 253	
Db	775 L-RABELLTVL 783	
RESULT 5		
ABB67112	ABB67112 standard; protein; 8805 AA.	
AC	ABB67112;	
XX		
XX	26-MAR-2002 (first entry)	
XX		
XX	Drosophila melanogaster polyprotein seq ID NO 28128.	
XX		
XX	Drosophila; developmental biology; cell signalling; insecticide;	
XX	pharmaceutical.	
XX		
XX	Drosophila melanogaster.	
XX		
XX	WO200171042-A2.	
XX		
XX	27-SEP-2001.	
XX		
XX	23-MAR-2001; 2001WO-US009231.	
XX		

PR 23-MAR-2000; 2000US-0191637P.
 PR 11-JUL-2000; 2000US-00614150.
 XX (PEKE) PE CORP NY.
 PA Venter JC, Adams M, Li PWD, Myers EW;
 PI MPI; 2001-656860/75.
 DR N-PSDB; AB111215.
 XX
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signaling and cell-cell
 PT interactions.
 XX
 PS Disclosure; SEQ ID NO 28128; 21pp + Sequence Listing; English.
 XX
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signaling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (AB116176-AB130511), expressed DNA
 CC sequences (AB101840-AB116175) and the encoded proteins (A8B57737-
 CC A8B72072). The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_sequences
 CC
 SQ Sequence 8805 AA;

Query Match 7.6%; Score 103; DB 4; Length 8805;
 Best Local Similarity 23.4%; Pred. No. 28;
 Matches 64; Conservative 39; Mismatches 108; Indels 62; Gaps 11;

QY 17 ABIELEAARTFKR--HAGRRVVDSDRGVTAAVST-----GRLIDYK----- 59
 DB 1474 SGIQFBEIRLKSGSQIGHSVDIAGIRPRTGRVLTTEAIGRLIDRTGEMLVGDR 1533
 QY 60 -----APTNGVIAHLRASKPLVLRVFTLSRNEIDVERGSKD--SDWEPVBEAAKQLA 112
 DB 1534 RITLEQAADGGLT-DIQLAKQLLEPGAGRDASGRELSLEVIQREISEAESGETAKRI 1592
 QY 113 FVEDRTTFEGYSAASIEGIRASASNPAULTPEPDRPREPDVTSQALSTRLAGVDGPRSYL 172
 DB 1593 KVNNITVTVE-----QTSGELGSPENPRNADAITAGSVDTK-----TGLYRVK 1635
 QY 173 LSAADVTKVSESDHGYPREH-----LNRIVDGDIIWAPAIQAFVLTTRGGDFDLQIG 227
 DB 1636 SGGTT--SLAEVREGTLRHESVTTKSNPLCSDAIHGLVDGAGWIDRNSGDKFRLD 1693
 QY 228 TDVA-----IGYASHDTPTERLYLOETL 250
 DB 1694 SAIANQLIDASVREVDKAKDT---KITLQEAL 1723

RESULT 6
 AAW72898
 ID AAW72898 standard; peptide; 19 AA.
 XX
 AC AAW72898;
 XX
 DT 21-JAN-1999 (first entry)
 XX
 DE Mycobacterium tuberculosis antigen CFP29 N-terminal peptide.
 XX
 KM Mycobacterium tuberculosis; antigen; vaccine; immunological; immunogen;
 KM infection.
 XX
 OS Mycobacterium tuberculosis.
 XX
 PN MO9844119-A1.
 XX
 PD 08-OCT-1998.
 XX

PF 01-APR-1998; 98WO-DK000132.
 XX
 PR 02-APR-1997; 97DK-00000376.
 PR 18-APR-1997; 97US-0044624P.
 PR 10-NOV-1997; 97DK-00001277.
 PR 05-JAN-1998; 98US-0070488P.
 XX
 PA (STAT-) STATENS SERUM INST.
 XX
 PI Andersen P, Nielsen R, Rosenkrands I, Weldingh K, Rasmussen PB;
 PI Oettinger T, Florio W;
 DR MPI; 1998-542705/46.
 XX
 PT New isolated mycobacteria polypeptides and nucleic acids - used for
 PT developing products for the diagnosis of or vaccination against
 PT mycobacterial infections, particularly tuberculosis.
 XX
 PS Claim 1; Page 144; 163pp; English.
 XX
 CC The present sequence represents a Mycobacterium tuberculosis N-terminal
 CC peptide. Products from the present invention, which describes protein
 CC fragments and nucleic acid fragments derived from M. tuberculosis, can be
 CC used in the detection of and prevention of mycobacterial infections. In
 CC particular, the proteins and nucleic acids can be used for the diagnosis
 CC of or vaccination against tuberculosis caused by M. tuberculosis, M.
 CC africanum or M. bovis
 CC
 SQ Sequence 19 AA;

Query Match 7.4%; Score 100; DB 2; Length 19;
 Best Local Similarity 100.0%; Pred. No. 0.0044;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNNTLYRDLA PYTEAAWAEI 19
 DB 1 MNNTLYRDLA PYTEAAWAEI 19

RESULT 7
 AAY21915
 ID AAY21915 standard; peptide; 19 AA.
 XX
 AC AAY21915;
 XX
 DT 06-SEP-1999 (first entry)
 XX
 DE N-terminal sequence of antigen CFP29.
 XX
 KM Immunogenic; Mycobacterium tuberculosis; immune response; infection;
 KM tuberculosis; fusion polypeptide; T-cell epitope; ESAT-6; MP759; TB;
 KM pharmaceutical; vaccination; M. africanum; M. bovis; CFP7A; CFP30A;
 KM CFP7B; CFP19; CFP27; CFP30A; RD1-ORF; CFP10A; CFP16; CFP19; CFP23;
 KM CFP25A; CFP30B; CFP7B.
 XX
 OS Mycobacterium tuberculosis.
 XX
 PN MO9924577-A1.
 XX
 PD 20-MAY-1999.
 XX
 PF 08-OCT-1998; 98WO-DK000438.
 XX
 PR 10-NOV-1997; 97DK-00001277.
 PR 05-JAN-1998; 98US-0070488P.
 PR 01-APR-1998; 98WO-DK000132.
 XX
 PA (STAT-) STATENS SERUM INST.
 XX
 PI Andersen P, Skjot R;
 XX
 DR MPI; 1999-347282/29.
 XX

PT New immunogenic fragment of Mycobacterium tuberculosis.

XX Example 3; Page 54; 265pp; English.

CC The invention describes a substantially pure immunogenic polypeptide
CC fragment (1) from Mycobacterium tuberculosis that is able to evoke a
CC protective immune response against infections by mycobacteria belonging
CC to the tuberculosis complex. The invention provides a (1) fusion
CC polypeptide comprising at least one polypeptide fragment (1) and at least
CC one fusion partner; (2) a fusion polypeptide fragment comprising a T-cell
CC epitope from M. tuberculosis protein BSAF-6, or MPT59 and a second
CC different amino acid sequence from M. tuberculosis, and/or including a
CC sequence which protects the first amino acid sequence from in vivo
CC degradation or post-translational processing; (3) a nucleic acid fragment
CC that encodes the above polypeptides. The polypeptides and nucleic acid
CC are useful as pharmaceuticals, for diagnosis of and as antigens for
CC vaccination against TB caused by Mycobacterium tuberculosis, africanum or
CC bovis. The polypeptides are also useful for diagnosing ongoing or
CC previous sensitization in an animal with bacteria belonging to the
CC tuberculosis complex. The invention also describes the use of CFP7A or
CC CFP30A or a T-cell epitope of for the induction of a strong immune
CC response in a mammal; use of CFP7B, CFP19 or MPT59-BSAF6 or a T-cell
CC epitope of for diagnosis of TB in a mammal by performing a DTH type skin
CC test; use of CFP27, CFP30A, RD1-ORF2, RD1-ORF3, RD1-ORF5, MPT59-BSAF6,
CC BSAF-MPT59, CFP10A, CFP16, CFP19, CFP23, CFP25A, CFP30B, CFP7B or a T-
CC cell epitope of for the preparation of an immunological composition; and
CC for the preparation of a subunit vaccine

SO Sequence 19 AA;

Query Match 7.4%; Score 100; DB 2; Length 19;
Best Local Similarity 100.0%; Pred. No. 0.0044;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNNTYRDLPVTEAAVAEI 19
DB 1 MNNTYRDLPVTEAAVAEI 19

RESULT 8

ABB91978 ABB91978 standard; protein; 755 AA.

AC ABB91978;

DT 31-MAY-2002 (first entry)

XX Herbicidally active polypeptide SEQ ID NO 1189.

KM Herbicidal; plant; agriculture; herbicide.

OS Arabidopsis thaliana.

PN WO200210210-A2.

PD 07-FEB-2002.

PF 28-AUG-2001; 2001WO-EP009892.

PR 28-AUG-2001; 2001WO-EP009892.

PA (FARB) BAYER AG.

PI Tietjen K, Weidner M;

DR WPI, 2002-269010/31.

PT Identifying plant target proteins for herbicidally active compounds,
PT comprising aligning and comparing nucleic acid or amino acid sequences
PT from plant with nucleic acid or amino acid sequences from non-plant
PT organisms.

PS Claim 5; SEQ ID NO 1189; 261pp + Sequence Listing; English.

XX The invention relates to identifying target proteins (ABB90790-ABB94016)

CC for herbicidally active compounds, comprising aligning and comparing

CC nucleic acid or amino acid sequences from plant with nucleic acid or

CC amino acid sequences from non-plant organisms using suitable search

CC parameters, where plant sequences having an E-value greater by a factor

CC of 3 than the E-value of most similar non-plant sequences are selected.

CC The polypeptides or nucleic acids encoding them are useful for

CC identifying modulators. The identified modulators are useful as

SO Sequence 755 AA;

Query Match 7.4%; Score 99.5; DB 5; Length 755;
Best Local Similarity 20.8%; Pred. No. 1.5;
Matches 71; Conservative 49; Mismatches 112; Indels 109; Gaps 18;

QY 11 VTBAWAEIELEAKRPFKRIAGR-----RVVDSDF--GGPVRAVSTGRLIDVKA 60

DB 68 ITIKKSPADY---KTYPERLDERVHELPVDMFVTYADVRBPPLIVNTVSLAVNY 124

QY 61 PTNGVIAHL-----RASK-----PL-----VRLRPPTLSRNEIDVBERG 95

DB 125 PANKLACTYSDDCSPITTSLSKASGFAKIWPFCCKKVVYRAAPMYTRNSPEAAGS 184

QY 96 SKSDWEPYVKAARKLA-FVEDRTIFEGYSA--SIEGIRASASNPALTLPE----- 144

DB 185 EFSKDWEMTKREYEKLSQKVEDATGSHWLDADBDFAFNTKSNDSHTIVKVENKGG 244

QY 145 --DPREIPDYI-----SQALSEL-RLAGV--DGPYSTLSADYTKVSET 184

DB 245 VGDEKEVPHVYISREKRPNHFHRYKAGANNFLVRVSGLMTNAPMYLAVDCDMVNEADV 304

QY 185 -----SDH---GYP-----IREHLRLVGDGIYMAPAIDGAF 213

DB 305 VROAMCTFLOKSDSNHCAVQYPODFYDSNGBELTTLQYLGRGNG--IQGFQYAGSG 362

QY 214 VLTTRGGDFDL---QGETDVAIGYASHDTTERTLYLQETLT 251

DB 363 CPHTRRVYVGLSLDDGDDSLSSIA---TRKYLABESLT 399

RESULT 9

ADN26519 ADN26519 standard; protein; 565 AA.

AC ADN26519;

DT 02-DEC-2004 (first entry)

XX Bacterial polypeptide #9172.

KM Recombinant DNA construct; transformed plant; improved plant property;

KM cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;

KM pathogen tolerance; pest tolerance; plant disease resistance;

KM cell cycle pathway modification; plant growth regulator;

KM homologous recombination; seed oil yield; protein yield; carbohydrate;

KM bacterial polypeptide.

OS Bacteria.

PN US2003233675-A1.

PD 18-DEC-2003.

PF 20-FEB-2003; 2003US-00369493.

PR 21-FEB-2002; 2002US-0360039P.

PA (CAOY/) CNO Y.
PA (HINK/) HINKLE G J.
PA (SLAT/) SLATER S C.

Db 127 GRVIDHLERGLDLSSEK---TLVLDADENKLRNGFTED---VEVLRLKLPASRQVALFS 180
Qy 122 GYSNASIEGIRASNSPALTLPEDPREIPDVISOALSELRLAGVDGPY---SVLLSADVY 178
Db 181 ATMPQRLRI-----AQYLDGPRI---VTIATTTTAANIRGRYMWVSGLHKDAL 229
Qy 179 TKVSETSDHGVPY-----REHLNRLVDGDIIMAPALDCAF-----VLTTRGDF 222
Db 230 TRILEVETFPDAMIIFVTKATESELAKLQARGLTAAINQDMQOQERRTIHQKDGKL 289
Qy 223 DLQGTVDVAIGYASHDPTREL 244
Db 290 DILVATDV---AARGLDVERI 307

RESULT 11
ADS28471
ID ADS28471 standard; protein; 614 AA.
XX
AC ADS28471;
XX
DT 02-DEC-2004 (first entry)
XX
DE Bacterial polypeptide #17504.
XX
KM Recombinant DNA construct; transformed plant; improved plant property;
KM cold tolerance; heat tolerance; drought tolerance; herbicide; osmotic;
KM pathogen tolerance; pest tolerance; plant disease resistance;
KM cell cycle pathway modification; plant growth regulator;
KM homologous recombination; seed oil yield; protein yield; carbohydrate;
KM nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
KM bacterial polypeptide.
XX
OS Bacteria.
XX
PN US2003233675-A1.
XX
PD 18-DEC-2003.
XX
PF 20-FEB-2003; 2003US-00369493.
XX
PR 21-FEB-2002; 2002US-0360039P.
XX
PA (CAOY/) CAO Y.
XX
PA (HINK/) HINKLE G J.
PA (SLAT/) SLATER S C.
PA (CHEN/) CHEN X.
PA (GOLD/) GOLDMAN B S.
XX
PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
XX
DR WPI; 2004-061375/06.
XX
PT New recombinant DNA construct comprising a promoter positioned to provide
PT for expression of a polynucleotide encoding a polypeptide from a
PT microbial source, useful for producing plants with improved properties.
XX
PS Claim 1; SEQ ID NO 17504; 122bp; English.

CC The invention relates to a recombinant DNA construct comprising a
CC promoter functional in a plant cell, where the promoter is positioned to
CC provide for expression of a polynucleotide encoding a polypeptide from a
CC microbial source. The invention also relates to a transformed plant
CC comprising the recombinant DNA construct and a method of producing a
CC transformed plant having an improved property. The plant is a crop plant
CC such as maize or soybean. The method of producing a transformed plant
CC having an improved property comprises transforming a plant with the
CC recombinant DNA construct and growing the transformed plant, where the
CC polynucleotide or polypeptide is useful for improving plant properties.
CC The recombinant DNA construct is useful for producing plants with
CC improved plant properties, e.g. improved cold, heat or drought tolerance,
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC increased resistance to plant disease, better growth rate by modification

CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition, improved lignin production or improved galactomannan
CC production. This sequence represents a bacterial polypeptide used in the
CC scope of the invention. Note: The sequence data for this patent did not
CC form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.
XX
SQ Sequence 614 AA;
Query Match 7.1%; Score 96; DB 8; Length 614;
Best Local Similarity 26.0%; Pred. No. 2.5;
Matches 68; Conservative 31; Mismatches 111; Indels 52; Gaps 13;
Qy 8 LAPVTEAMABIEHAKRTKRH---TAGRRVYDV---SDPGGPVTAAVSGRLIDVCAPT 62
Db 89 LAPVTELA-----IQVAEAFQRYAASISGFRVLPVVGQSGYGOQLALKRGVHVIGTP- 142
Qy 63 NGVIAHL-RASKPLVRLRVPEFTLSRNEIDDERGSKDSMEBPVGEAKKLAIVEDRTIFE 121
Db 143 GRVIDHLERGLDLSSEK---TLVLDADENKLRNGFTED---VEVLRLKLPASRQVALFS 196
Qy 122 GYSNASIEGIRASNSPALTLPEDPREIPDVISOALSELRLAGVDGPY---SVLLSADVY 178
Db 197 ATMPQRLRI-----AQYLDGPRI---VTIATTTTAANIRGRYMWVSGLHKDAL 245
Qy 179 TKVSETSDHGVPY-----REHLNRLVDGDIIMAPALDCAF-----VLTTRGDF 222
Db 246 TRILEVETFPDAMIIFVTKATESELAKLQARGLTAAINQDMQOQERRTIHQKDGKL 305
Qy 223 DLQGTVDVAIGYASHDPTREL 244
Db 306 DILVATDV---AARGLDVERI 323

RESULT 12
ABB66212
ID ABB66212 standard; protein; 621 AA.
XX
AC ABB66212;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster polypeptide SEQ ID NO 25428.
XX
KM Drosophila; developmental biology; cell signalling; insecticide;
KM pharmaceutical.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US009231.
XX
PR 23-MAR-2000; 2000US-0191637P.
PR 11-JUL-2000; 2000US-00614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
DR WPI; 2001-656860/75.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions.

PS Disclosure; SEQ ID NO 25428; 21pp + Sequence Listing; English.
 XX
 XX The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from *Drosophila*. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (AB116176-AB130511), expressed DNA
 CC sequences (AB101840-AB116175) and the encoded proteins (AB57737-
 CC AB872072). The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at http://wipo.int/pdb/published_pcr_sequences
 CC
 XX
 SQ Sequence 621 AA;
 Query Match 7.1%; Score 96; DB 4; Length 621;
 Best Local Similarity 22.6%; Pred. No. 2.5; 109; Indels 58; Gaps 12;
 Matches 61; Conservative 42; Mismatches 109; Indels 58; Gaps 12;
 QY 17 ARIELEAARTFKKHIA-----GRRVVDVSDPGP-----VTAVSTGRLLDYKAPT 62
 DB 110 ADVAAEAGGTTGHIAGFTVLTNGSRVTFLDTPGHAAFSAMBARGVATDIIIVVAAH 169
 QY 63 NGYIALRLRASKPLVR-LRVPTLSRNEIDVERGSKDSMEPVYKAACKLAFFEDR---- 117
 DB 170 DGVAQTRREVTLQAKAQVPIIIVALKNKIDKPEANIEKSK-----RELAQMGLEHGGDV 225
 QY 118 --TIFPGYSAAISIEGIRASASNPAA--LTPEDEPREIPD-VISQALSELRLA----- 163
 DB 226 QVPIPALKGTNLELAEAVSTQATLMKADPTGLVEGIVGSEKTDPPRGKLSIAVSR 285
 QY 164 GVDGPPSVLLSADVTVKSETSDH-GYPIREHLNRLVDGDIWAPAIDGAFVLTTR---- 218
 DB 286 GTLRKQSVLLSGLAHAKVNGLPHNGQPLSE-----APGTPVEILGWMLPL 313
 QY 219 GGDPLQLQGLTD---VAIGYASHDTTERL 244
 DB 334 AGDLILEVETEKKAHAVALKTYREHESQOEKI 363
 RESULT 13
 AAG99871 ID AAG99871 standard; protein; 4038 AA.
 XX
 XX AAG99871;
 DT 16-JUN-2003 (first entry)
 XX
 XX S. cinamomensis MonAIV/polyketide synthase multi-enzyme MONS4.
 DE
 XX
 XX Monensin; gene cluster; polyketide synthase; antibiotic; antihelminthic;
 KM insecticide; immunosuppressant; antifungal; antibacterial; polyether;
 KM mon BI; mon BII; mon CI; mon CII; mon H; mon RI; mon RII; mon T; mon AIX;
 KM mon AX.
 KW
 XX Streptomyces cinamomensis.
 OS
 XX WO200168867-A1.
 PN
 XX 20-SEP-2001.
 PD
 XX 30-MAY-2000; 2000WO-GB002072.
 PF
 XX 28-MAY-1999; 99GB-00012563.
 PR
 XX (BIOT-) BIOTICA TECHNOLOGY LTD.
 PA
 XX Leadlay PF, Staunton J, Ollmyk M;
 PI
 XX WPI; 2001-611393/70.
 DR
 XX N-PSDB; ABX04971.
 DR
 XX New DNA sequence encoding polyketide synthase, useful for the production

PT of polyketides such as antibiotic monensin.
 XX
 XX Claim 6; Page 82-85; 212pp; English.
 XX
 XX The invention relates to a DNA sequence which is a fully defined sequence
 CC of 103551 base pairs appearing as ABX04971, or its variant, that it is
 CC not a sequence encoding all or part amino acids 1-920 encoded by mon AI
 CC as given in the specification. The DNA is the S. cinamomensis polyketide
 CC antibiotic monensin biosynthetic gene cluster. Also included are a
 CC recombinant cloning or expression vector comprising the gene cluster, a
 CC transformant host cell which has been transformed to contain the gene
 CC cluster (and is capable of expressing a corresponding polypeptide), a
 CC hybridization probe derived from the gene cluster (for identification and
 CC isolation of the same or analogous gene cluster, e.g. one which binds
 CC specifically to a region of the monensin gene cluster selected from mon
 CC BI, mon BII, mon CI, mon CII, mon H, mon RI, mon RII, mon T, mon AIX and
 CC mon AX), the use of the mon RI gene or variant and a monensin promoter to
 CC control expression of a heterologous gene in *Streptomyces cinamomensis*,
 CC a polypeptide encoded by a portion of the monensin gene cluster
 CC (preferably comprising mon BI, mon BII, mon AIX or mon AX or their
 CC mutants, alleles or variants), an epoxide enzyme encoded by mon CI, a
 CC cyclase enzyme encoded by mon CII, producing S. cinamomensis capable of
 CC enhanced levels of production of monensin comprising engineering it to
 CC overexpress the mon RI gene, S. cinamomensis containing multiple copies
 CC of the mon RI gene and/or its variants, expressing a gene heterologous to
 CC S. cinamomensis comprising transforming S. cinamomensis with DNA
 CC encoding a heterologous gene and expressing the gene under control of the
 CC activator gene mon RI or acII/orf4 and 13-propyl erythromycin A. The
 CC processes and materials (enzyme systems, nucleic acids and vectors) are
 CC useful for preparing polyketides by recombinant synthesis. The
 CC polyketides are useful as insecticides, antibiotics, antihelminthics,
 CC antifungals, antibacterials or other pharmaceuticals. In particular the
 CC gene is useful for the production of monensin, an antibiotic polyether
 CC polyketide. The present sequence represents a protein encoded by the
 CC monensin gene cluster
 CC
 XX
 SQ Sequence 4038 AA;
 Query Match 7.0%; Score 95; DB 4; Length 4038;
 Best Local Similarity 25.3%; Pred. No. 57;
 Matches 60; Conservative 29; Mismatches 86; Indels 62; Gaps 14;
 QY 8 LAPVTEA-----AAAEIELEAARTFKKHIA-----AGRRVVDVSDPGPVTAAVSTGRLL 56
 DB 1176 LIPPTDAEQVWLPFAANNVDALNAHVATVVRVLTPLGRI-----DGLRIIVADAVG--- 1228
 QY 57 DVKAPTNVIAHLRASKPLVRLRVPTLSRNEIDVERGSKDSMEPVYKAACKLAFFEDR 116
 DB 1229 --APVL-TVRDLR-SRPDTGRIAAATRD-----RGLFLDEWIAPEMAENNA----- 1274
 QY 117 RTIFEGYSAAISIEGIRASASNPALTLPEDPREIPDVISQALSELRLAGVDPYSVLLSAD 176
 DB 1275 -----GPRDASSEGWVTLGEDAASLADLNASV-----EAGAPAPQ--LVAAIP 1314
 QY 177 VTTKVSSETSDHGYPIREHLNRLVDGDIW--APAIAGAVLTTGGDDPLQLQGLTDVA 231
 DB 1315 V--BPDRTDDGIALATHVLDLVQ--TWIASPLHDSRLVLTVRGAVTDAV--DVA 1363
 RESULT 14
 AAB26917 ID AAB26917 standard; protein; 853 AA.
 XX
 XX AAB26917;
 DT 12-JUN-2001 (first entry)
 XX
 XX Large subunit of periplasmic nitrate reductase NapA.
 DE
 XX
 XX NapA; periplasmic nitrate reductase; enzyme; transgenic organism;
 KM toxic waste denitrification; denitrifying bacteria.
 KM
 XX Pseudomonas sp. G-179.
 OS

XX US6136588-A.
 XX 24-OCT-2000.
 PD
 XX
 PF 15-JUL-1999; 99US-00354129.
 XX
 PR 17-JUL-1998; 98US-0093191P.
 XX
 PA (DUPO) DU PONT DE NEMOURS & CO E. I.
 XX
 PI Bedzyk LA, Ye RW;
 XX
 DR WPI; 2000-686079/67.
 DR N-PSDB; AAA94687.
 XX
 PT New nucleic acid encoding bacterial nitric oxide reductase, useful e.g.
 PT for preparing transgenic organisms able to detoxify wastes.
 XX
 PS Example 2; Col 35-40; 24pp; English.
 XX
 CC The present sequence is Napa from *Pseudomonas* sp. strain G-179. This
 CC sequence is the large subunit of periplasmic nitrate reductase. The Napa
 CC gene can be used to create transgenic organisms that can denitrify toxic
 CC waste. In addition, the present sequence may be used to produce probes
 CC and primers for identification of new denitrifying bacteria. Residues 20
 CC to 853 of the present sequence are encoded by the sequence in AAA94687
 XX
 SO Sequence 853 AA;

Query Match 7.0%; Score 94.5; DB 3; Length 853;
 Best Local Similarity 25.1%; Pred. No. 5.8;
 Matches 59; Conservative 27; Mismatches 74; Indels 75; Gaps 15;
 QY 77 RLKVPFTLSRNEIDVVRGSKDSMDMPV-----KEAKKLAFFVDR----- 117
 DB 119 RLTPPLLRKNGVFD-----KEGFEPVTEBAPDIAERAKKT--LKEKPTALGMFGS 171
 QY 118 ---TIFEGYASASI--EGIRASASNPALTPEDEPREIPDIVISQALSELRLAGVDP---Y 169
 DB 172 GQRTIFEGYAAITKLMRAGFRSNL-----DPAHCHMASAAYGFRKRTGMBEPMGCY 223
 QY 170 SVLLSADVTKVSETSDHGYPIREHNLRLVGDIIWAPADGAF-----VLT---R 218
 DB 224 DDFEHADAFLMGSNMAEMHP-----ILWTRLADRLRGHEHVKSVLSTFTHR 271
 QY 219 GGPB-DLQI---GTDAI-GY-ASHDTBERL---YLQETTLFLCTTAASVAL 263
 DB 272 SMDLADIPLVFKPGTDIAIINLYIANHIQTGRVQVEFIDKTKEMQATTIDIGYL 326

RESULT 15
 ABU62618
 ID ABU62618 standard; protein; 853 AA.
 AC
 XX ABU62618;
 XX
 DT 23-OCT-2003 (revised)
 DT 02-SEP-2003 (first entry)
 XX
 DE *Pseudomonas* nitrate reductase subunit A, Napa.
 XX
 KM Periplasmic nitrate reductase subunit; enzyme; Napa; NapB; NapC;
 KM denitrification; water purification; industrial waste purification;
 KM denitrifying bacteria; nitric oxide reductase subunit; NorB; NorC; NorQ;
 NorD.
 OS
 XX *Pseudomonas* sp; strain G-179.
 XX
 PN US6429003-B1.
 XX
 PD 06-AUG-2002.
 XX

PF 15-FEB-2000; 2000US-00504357.
 XX
 XX 17-JUL-1998; 98US-0093191P.
 PR 15-JUL-1999; 99US-00354129.
 XX
 PA (DUPO) DU PONT DE NEMOURS & CO E. I.
 XX
 XX Bedzyk LA, Ye RW;
 XX
 DR WPI; 2003-521168/49.
 DR N-PSDB; ACD26043.
 XX
 PT New *Pseudomonas* bacterial periplasmic nitrate reductase subunit nucleic
 PT acid, for producing enzymes used in water purification.
 XX
 PS Claim 1; Col 35-40; 24pp; English.
 XX
 CC The invention relates to an isolated nucleic acid fragment encoding an
 CC enzymatically active bacterial periplasmic nitrate reductase subunit from
 CC *Pseudomonas* sp. strain G-179, which comprises NapA, NapB and NapC
 CC subunits. Also included are a chimeric gene comprising the Nap nucleic
 CC acid linked to regulatory sequences, a transformed cell comprising the
 CC chimeric gene and a nucleic acid fragment encoding an enzymatically
 CC active bacterial periplasmic nitrate reductase subunit, obtained by
 CC probing a library with the Nap nucleic acid and sequencing the isolated
 CC clone. The Nap nucleic acid is useful for producing enzymes useful in
 CC denitrifying reactions (for water purification and industrial waste
 CC purifying) and for identifying other denitrifying bacteria. Also
 CC disclosed are the nucleic acid and protein sequences for the nitric oxide
 CC reductase subunit genes NorB, NorC, NorQ and NorD. The present sequence
 CC represents the Napa protein. Note: The present sequence is stated to be
 CC encoded by the DNA appearing as ACD26043 but is not encoded by it.
 CC (Updated on 23-OCT-2003 to standardise OS field)
 XX
 SO Sequence 853 AA;

Query Match 7.0%; Score 94.5; DB 7; Length 853;
 Best Local Similarity 25.1%; Pred. No. 5.8;
 Matches 59; Conservative 27; Mismatches 74; Indels 75; Gaps 15;
 QY 77 RLKVPFTLSRNEIDVVRGSKDSMDMPV-----KEAKKLAFFVDR----- 117
 DB 119 RLTPPLLRKNGVFD-----KEGFEPVTEBAPDIAERAKKT--LKEKPTALGMFGS 171
 QY 118 ---TIFEGYASASI--EGIRASASNPALTPEDEPREIPDIVISQALSELRLAGVDP---Y 169
 DB 172 GQRTIFEGYAAITKLMRAGFRSNL-----DPAHCHMASAAYGFRKRTGMBEPMGCY 223
 QY 170 SVLLSADVTKVSETSDHGYPIREHNLRLVGDIIWAPADGAF-----VLT---R 218
 DB 224 DDFEHADAFLMGSNMAEMHP-----ILWTRLADRLRGHEHVKSVLSTFTHR 271
 QY 219 GGPB-DLQI---GTDAI-GY-ASHDTBERL---YLQETTLFLCTTAASVAL 263
 DB 272 SMDLADIPLVFKPGTDIAIINLYIANHIQTGRVQVEFIDKTKEMQATTIDIGYL 326

Search completed: April 14, 2006, 17:18:26
 Job time : 106.677 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: April 14, 2006, 17:19:08 ; Search time 15.539 Seconds

(without alignments)
1640.866 Million cell updates/sec

Title: US-10-620-246-16

Perfect score: 1348

Sequence: 1 MNNLYRDLPVTEBAWAEIE.....LQETLFLCYTAASVALSH 265

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

SUMMARIES

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Query Match	Length	ID	Description
1	1341	99.5	265	2 E70520	hypothetical prote
2	384	28.5	262	2 B72333	bacteriocin - Ther
3	142	10.5	281	2 E70451	hypothetical prote
4	111	8.2	492	1 S68968	trypanothione-disu
5	105.5	7.8	390	2 B72639	hypothetical prote
6	105.5	7.8	467	2 E75395	conserved hypochet
7	102.5	7.6	834	2 B95346	NADa periplasmic n
8	100	7.4	694	2 B84331	hypothetical prote
9	99.5	7.4	755	2 T02553	cellulose synthase
10	96.5	7.2	524	2 AB2338	acetyl-coenzyme A
11	96.5	7.2	1489	2 S73015	polypeptide synthas
12	96	7.1	464	2 T50734	regulatory protein
13	96	6.9	614	2 G82630	ATP-dependent RNA
14	93.5	6.9	673	2 A45456	NADH2 dehydrogenas
15	92	6.8	255	2 B70982	probable nei prote
16	91.5	6.8	418	2 G70350	conserved hypochet
17	91	6.8	469	2 A84578	probable WD-40 rep
18	90.5	6.7	463	2 C86034	L-seryl-CRNasec se
19	90.5	6.7	463	2 D91187	L-seryl-CRNasec se
20	90.5	6.7	463	2 A65159	L-seryl-CRNasec se
21	90	6.7	331	2 G82919	fatty acid/phospho
22	90	6.7	346	2 D90450	conserved hypochet
23	90	6.7	372	2 F83156	probable metal tra
24	89.5	6.6	377	2 E75388	conserved hypochet
25	89.5	6.6	539	2 F70737	chaperonin groEL1
26	89.5	6.6	566	2 A75481	sensor histidine k
27	89.5	6.6	2314	2 T28688	hypothetical prote
28	89	6.6	309	2 F75283	hydroxymethylbilan
29	89	6.6	570	2 D72597	hypothetical prote

30	89	6.6	1772	2 T36105	probable large gly
31	88.5	6.6	281	2 AE3523	icc protein (impor
32	88.5	6.6	3149	1 Q0B88	BPEFI protein - hu
33	88	6.5	457	2 P97533	dihydroliipoamide a
34	88	6.5	6831	2 A88852	protein unc-22 (im
35	88	6.5	6839	2 S57242	twichin (similar
36	88	6.5	7160	2 T27935	hypothetical prote
37	87.5	6.5	371	2 T50541	riboflavin-specifi
38	87.5	6.5	507	2 C96973	spore germination
39	87.5	6.5	729	2 A69202	conserved hypochet
40	87.5	6.5	742	2 H83529	nitrate reductase
41	87.5	6.5	831	2 S50153	endopeptidase Ctp
42	87	6.5	908	1 S61476	hypothetical prote
43	86.5	6.5	1220	2 T32916	probable metal-tra
44	86.5	6.4	740	2 C83185	ribonucleotide red
45	86.5	6.4	937	2 D87483	

ALIGNMENTS

RESULT 1

E70520 hypothetical protein Rv0798c - Mycobacterium tuberculosis (strain H37Rv)

C:Species: Mycobacterium tuberculosis

C>Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004

C/Accession: E70520

R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, J.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.

Rajandream, M.A.; Rogers, J.; Kuter, S.; Seeger, K.; Skelton, S.; Squares, S.

Nature 393, 537-544, 1998

A:Authors: Squares, R.; Sultson, J.E.; Taylor, K.; Whitehead, S.; Barrett, B.G.

A>Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome

A:Reference number: A70500; PMID:98295987; PMID:9634230

A:Accession: E70520

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-265 <COL>

A:Cross-references: UNIPROT:O07181; UNIPARC:UP100000D5E0F; GB:Z96797; GB:AL123456; NID:

A:Experimental source: strain H37Rv

C:Genetic81

A:Gene: Rv0798c

Query Match 99.5%; Score 1341; DB 2; Length 265;
Best Local Similarity 99.6%; Pred. No. 2e-98;
Matches 264; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY	1	MNNLYRDLPVTEBAWAEIELEAARTPKRIAGRVVDS	PGCPVTAAVSTGRLIVKA	60
DB	1	MNNLYRDLPVTEBAWAEIELEAARTPKRIAGRVVDS	PGCPVTAAVSTGRLIVKA	60
QY	61	PTNGVTAHAPASKPVLAIRPFTLSRNEIDVRSKDS	QSWPEYKAKKLAIPEDRTIF	120
DB	61	PTNGVTAHAPASKPVLAIRPFTLSRNEIDVRSKDS	QSWPEYKAKKLAIPEDRTIF	120
QY	121	EGYSAASIEGIRASNSPALTLPEDPREIPDIVISQ	LSRLAGVDPYVLSADVYTK	180
DB	121	EGYSAASIEGIRASNSPALTLPEDPREIPDIVISQ	LSRLAGVDPYVLSADVYTK	180
QY	181	VSESTDHGYPRIHRLNRLVVDGIIMAPADIGAFVL	TRGSDFDLQGTGVAIGYASHDT	240
DB	181	VSESTDHGYPRIHRLNRLVVDGIIMAPADIGAFVL	TRGSDFDLQGTGVAIGYASHDT	240
QY	241	TERLYLOETLFLCYTAASVALSH	265	
DB	241	TERLYLOETLFLCYTAASVALSH	265	

RESULT 2

B72333 bacteriocin - Thermotoga maritima (strain MSB8)

C:Species: Thermotoga maritima

C>Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 09-Jul-2004

C/Accession: B72333
 R/Release: K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.D.; Haft, D.H.; Hickey
 Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.;
 C.M.
 Nature 399, 323-329, 1999
 A/Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome seq
 A/Reference number: A72200; MUID:99287316; PMID:10360571
 A/Accession: B72333
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-262 <ARN>
 A/Cross-references: UNIPROT:Q9WZP2; UNIPARC:UP10000007358; GB:AE001747; GB:AE000512; NID:
 A/Experimental source: strain MSB8
 C/Genetics: TM0785
 A/Gene: TW0785

Query Match 28.5%; Score 384; DB 2; Length 262;
 Best Local Similarity 36.2%; Pred. No. 6.2e-23;
 Matches 92; Conservative 48; Mismatches 108; Indels 6; Gaps 6;

QY 1 MNLYDLAVTEAAMAEIELEAARTPKRHIAGRVVDSDPGPVTAAVSTGRLIDYKA 60
 DB 4 MEPLKSPAPLTERKQWEIDNRAREIKTQYGRKFDVGGPYGWEYAAHPLGE-VEVLS 62
 61 PTNGVIA-HLRASKPLVRLRPPTLSRNEIDVERGSKSDWBPVKAARKLAFFVEDRTI 119
 DB 63 DEBEVVMGRLKSLPILELATFTLDMEIDNLERKPNVDLSLEBTVAKVAEFEDVI 122
 QY 120 PEGYSAASIGCISASSNPALTLPEDREIPDVYSQSLSELRLAGVDPYSLADVYT 119
 DB 123 PRGCEKSGVGLISFEER-KIECGSTPKDLALVRLSIFSKDGEPTVLTINTRWI 181
 QY 180 K-VSETSDHGYPRREHNR-LVDDDIWAPIDAFAFLTRRGSGDFDLQGTDAVIGASH 237
 DB 182 NPLKEBAGH-YPLAKRYEBCLRGSKITTRIEDALVVSERGDFKILQDLSIGEDR 240
 QY 238 DTDTERTLYLOETLT 251
 DB 241 EKDAVRLPIETETPT 254

RESULT 3
 E70451
 hypothetical protein aq_1760 - Aquifex aeolicus
 C/Species: Aquifex aeolicus
 C/Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 09-Jul-2004
 C/Accession: E70451
 R/Recker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; O'
 V.
 Nature 392, 353-358, 1998
 A/Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
 A/Reference number: A70300; MUID:9819666; PMID:9537320
 A/Accession: E70451
 A/Status: preliminary; nucleic acid sequence not shown; translation not shown
 A/Molecule type: DNA
 A/Residues: 1-281 <AQF>
 A/Cross-references: UNIPROT:O67639; UNIPARC:UP100000566DE; GB:AE000754; NID:92984047; PI
 A/Experimental source: strain VFS
 C/Genetics:
 A/Gene: aq_1760

Query Match 10.5%; Score 142; DB 2; Length 281;
 Best Local Similarity 24.2%; Pred. No. 0.00084;
 Matches 63; Conservative 41; Mismatches 120; Indels 36; Gaps 9;

QY 1 MNLYDLAVTEAAMAEIELEAARTPKRHIAGRVVDSDPG- - - - -P 45
 DB 1 MEFLORDQAPLTAEMFQIDKTAYEVFKSTVCRKFPVVGPFQAGQVVSVDLYGVEP 60
 QY 46 VTAAVSGRLIDYKAP-NGVIALRLASKPLVRLRPPTLSRNEIDVERGSKSDWBPY 104
 DB 61 GVCENVKGOEYKVCPEVTRGERKIV- - - - -PVPTLYKDFVLSWRDLERHROFNLPIVDITGV 116

QY 105 KEAAKLAFAVEDRTIEGYSASISIEGIRSSNPALTLPEDPRE- - - - -IPDVISQAL 157
 DB 117 AAAASLAAVEDKLIIFGNQEMGIEGFLPAKG- - - - -TLREKLSDEWKEGNAPQDVV-KGI 171
 QY 158 SEIRLAGVDPYVLSADVYTTVSTSDH-GVPIREHNRIVDGIINAPAI--DGAFV 214
 DB 172 SRLVEKGYFTNYTLINPKRYFLLNRIDHNTGLLEIQIKVKV-EVYQTPILIPDVLVL 230
 QY 215 LTRRGDFDLQGTDAVIGY 234
 DB 231 VSASPANFDALMDVNVAF 250

RESULT 4
 S68968
 trypanothione-disulfide reductase (EC 1.8.1.12) - Trypanosoma cruzi

C/Species: Trypanosoma cruzi
 C/Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
 C/Accession: S68968; S24243; S30204
 R/Borges, A.; Cunningham, M.L.; Tovar, J.; Fairlamb, A.H.
 Eur. J. Biochem. 228, 745-752, 1995
 A/Title: Site-directed mutagenesis of the redox-active cysteines of Trypanosoma cruzi t
 A/Reference number: S68968; MUID:95255281; PMID:7737173
 A/Accession: S68968
 A/Status: preliminary; nucleic acid sequence not shown; translation not shown
 A/Molecule type: DNA
 A/Residues: 1-492 <BOR>
 A/Cross-references: UNIPROT:P28593; UNIPARC:UP1000016C0BD; EMBL:Z13958; NID:9624037; PI
 A/Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1995
 R/Borges, A.; Fairlamb, A.
 submitted to the EMBL Data Library, June 1992
 A/Description: Mutagenesis of the redox-active cysteines in the trypanothione reductase
 A/Reference number: S24243
 A/Accession: S24243
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-111, 'E', 113-139, 'A', 141-155, 'N', 157-492 <BOR>
 A/Cross-references: UNIPARC:UP100001720C2; EMBL:Z13958
 R/Sullivan, P.X.; Walsh, C.T.
 Mol. Biochem. Parasitol. 44, 145-148, 1991
 A/Title: Cloning, sequencing, overproduction and purification of trypanothione reductas
 A/Reference number: S30204; MUID:9187059; PMID:2011150
 A/Accession: S30204
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-94, 'K', 96-111, 'E', 113-155, 'N', 157-352, 'N', 354-401, 'NI', 404-440, 'V', 442-49
 A/Cross-references: UNIPARC:UP1000013788B; EMBL:M80851; NID:9162316; PIDN:AA63547.1; P
 C/Superfamily: dihydrolipoamide dehydrogenase; dihydrolipoamide dehydrogenase homology
 C/Keyword: NADP; oxidoreductase; redox-active disulfide
 F;9-468/Domain: dihydrolipoamide dehydrogenase homology <DLD>
 F;53-58/Disulfide bonds: redox-active #status predicted

Query Match 8.2%; Score 111; DB 1; Length 492;
 Best Local Similarity 25.8%; Pred. No. 0.52;
 Matches 76; Conservative 41; Mismatches 96; Indels 82; Gaps 18;

QY 13 EAAMAEIELEAARTPKRHIAGRVVDSDPGVTAAGTGLIDV- - - - -KAPTGV- 65
 DB 19 EAMW- - - - -NAATLYKKRAV- - - - -VIDQVHGSGPPPSALGCTGVNCGVPKLMTVGAQY 70
 QY 66 IAHLRAS- - - - -KPLVRLRPPTLSRNE-IDVER- - - - -GSK 97
 DB 71 MEHLRSASAGWGFDRITTLRAEKNLIVAMDVAVLNINKSYDMPDTEGLAEPFLAGSL 130
 QY 98 DS-DWEPYKAAKLAFAVERKRTIFEGYSAS- - - - -IEGRSA-SSNPALTLPEDPR 147
 DB 131 BSKNVNVRESADPAVAVERKLETHIILASGSGVPHMPNIPGIEHCISSEAAVYLPBPPR 190
 QY 148 EIPDVISQALSELQGV- - - - -DGYVLSADVYTTVSTSDHGYPIREHNRIVLD 200
 DB 191 RVLTVGGRFS-VERPACIFPAAYPKDQGVTLCTRGEM- - - - -ILRGFDH- - - - -TLREELTKQLT 244
 QY 201 GDIIWA- - - - -IDGAFVLTTRGG- - - - -DFDLQGTDAVIGASHDTDTTE 242

Db 245 ANGIQILTKENPAKVELINADSGSKVTEPSGRKDFDLMV---VAIGSPRTKDLQ 296

RESULT 5

B72639 hypothetical protein APE0546 - Aeropyrum pernix (strain K1)

C/Species: Aeropyrum pernix

C/Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004

C/Accession: B72639

R;Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Halkawa, Y.; Jin-no, K.; Takahawa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K DNA Res. 6, 83-101, 1999

A/Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyrum

A/Reference number: A72450; MUID:9310339; PMID:10382966

A/Accession: B72639

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-390 <RAW>

A/Cross-references: UNIPROT:Q9YEN2; UNIPARC:UP1000005DBB7; DDBJ:AP000060; NID:G5104188;

A/Experimental source: strain K1

C/Genetics:

A/Gene: APE0546

C/Superfamily: Methanococcus jannaschii conserved hypothetical protein M0041

Query Match

Best Local Similarity 7.8%; Score 105.5; DB 2; Length 390;

Matches 57; Conservative 34; Mismatches 87; Indels 55; Gaps 14;

Db 11 VTEAAMAEILEAARTFKRIHAGRRVVDSPGGPVTAAVSTGR--LIDYKAPITNGVI-- 66

Db 181 VEOALMPVLKLTGGERLVVAAGREDVD-----ARMGSGRPMIVGVSPRRRIPL 232

Qy 67 -----AHLRASKDPLRLRVFPTLSRNEI---DDVERGSKDSDPEPYKAAKGLAFVDR 117

Db 233 ERIEBAANAGCKGLVRRFPTPAKRAVALYKESTAR-----VRKYRRLVAAGE- 282

Qy 118 TIFEGYSAAISIEGIRASNSNPALTLPPDPREI---PDVI--SQMLSELRAGVGP--YS 170

Db 283 -----GVSEVVDVGRRLRLEG-AVIMQRTPSRVLHRRDILRRRLVLDGSPLEGARLME 337

Qy 171 VLISAD--VTYKVSSTSDHGY---PIREHLNR---LVGDGIIV---APAIDG 211

Db 338 CILEAGGLYIKELVSGDGRTRPSFAVLGRVAVCIELDVVWEHAPAPAG 390

RESULT 6

B75395

conserved hypothetical protein - Deinococcus radiodurans (strain R1)

C/Species: Deinococcus radiodurans

C/Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 05-Oct-2004

C/Accession: B75395

R;White, O.; Eiben, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.; M.; Shen, M.; Vamathavan, J.J.; Lam, P.; McDonald, L.; Uterback, T.; Zalewski, C.; Ma

S.; Smith, H.O.; Venter, J.C.; Fraser, C.M. Science 286, 1571-1577, 1999

A/Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.

A/Reference number: A75250; MUID:20036896; PMID:10567266

A/Accession: B75395

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-467 <NHI>

A/Cross-references: UNIPROT:Q9RUP1; UNIPARC:UP1000013AAC4; GB:AE001988; GB:AE000513; NID

A/Experimental source: strain R1

C/Genetics:

A/Gene: DR1435

A/Map position: 1

C/Superfamily: uncharacterized conserved protein

Query Match

Best Local Similarity 7.8%; Score 105.5; DB 2; Length 467;

Matches 80; Conservative 25; Mismatches 87; Indels 115; Gaps 18;

Qy 21 LEAARTFKRIHAGRRVVDSPGGPVTAAVSTGR---IDYKAP--TNGVIAHLRASKPL 75

Db 161 LTGKTHSSNITA--VTVAADDG3-----SSGRLREALMVAFGDLTDCYALSSPAL 212

Qy 76 VLRLVP-----FTLSRNEIDVERGSKDSDPEPYKAAKGLAFV--- 114

Db 213 ARLLHRRFGEGLEGHTFGNLLATYSR-----EKGIGCTMDDHLELTKRGRRVPA 266

Qy 115 -----BDRIT-----PEGSAASIGIRASNSNP-ALT-LPPDPRE----- 148

Db 267 TTRPVTVAELADGRTRTGRSRRFARQRPRIIRVRLPENPSALITQVLAVDAEMIVL 326

Qy 149 -----IDPVISQALSELRLAGVDGPYSVLADVYTKVSTF-----SDHGY 189

Db 327 GPGSLFTSIIPALLIPD-IRAAVR-----SPAPVYVASIMTEBERGTDGLSDHVN 378

Qy 190 PIREHLNRVLVDGIIIV-----PAI-----DGAFLVTTTGDDPL-----QLGIDVA 231

Db 379 AITRHIGRTPD-----WLLSNSKIEPAVQRRYQOEGATVTLDGAGRLGRVAFAPLIQ 434

Qy 232 IGYASHD 238

Db 435 AGTRARD 441

RESULT 7

B95346

Napa periplasmic nitrate reductase [imported] - Sinorhizobium meliloti (strain 1021) mag

C/Species: Sinorhizobium meliloti

C/Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 09-Jul-2004

C/Accession: B95346

R;Barnett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; Bower

.; Kalman, S.; Keating, D.H.; Palm, C.; Beck, M.C.; Surzycki, R.; Wells, D.H.; Yeh, K.C

Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001

A/Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium meliloti

A/Reference number: A95262; MUID:2139509; PMID:11481432

A/Accession: B95346

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-834 <NUR>

A/Cross-references: UNIPROT:Q92Z36; UNIPARC:UP100000CB17C; GB:AE006469; PIDN:AAK5332.1

A/Experimental source: strain 1021, megaplasmid pSymA

R;Galibert, F.; Finam, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler

pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.

L.; Hyman, R.W.; Jones, T. Science 293, 668-672, 2001

A/Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; LeLaure

hebalet, P.; Vandenberg, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yen, K

A/Title: The composite genome of the legume symbiont Sinorhizobium meliloti.

A/Reference number: A96039; MUID:21368234; PMID:11474104

A/Contents: annotation

C/Genetics:

A/Gene: napa

A/Genome: plasmid

C/Superfamily: Alcaligenes eutrophus nitrate reductase A chain

Query Match

Best Local Similarity 7.6%; Score 102.5; DB 2; Length 834;

Matches 65; Conservative 37; Mismatches 75; Indels 99; Gaps 16;

Qy 43 GGPVTAAGTGRIT---DVKAPPTNGVIAHLRA---SKPLV---RLRVPTLSRNEIDV 92

Db 56 GCGVMGVGEGGVVAHGHGMQAVNGINCINKYFYSKIMYGRDLKTPLLRRGN----- 110

Qy 93 ERGSKDSDPEPVK-----EAAKGLAFVDR-----TIFEGYSAASI-- 128

Db 111 GAFAKDGEEFVPSWDEAFVMAEQAKV--LKKDKGPTAVGMFGSGQMTTFEGYAAATLKR 168

Qy 129 EGIRASNSNPALTLPPDPREIPDVISQALSELRLAGVDGP---YSVLASDVYTKVSETS 185

Db 169 AGRSRNNL-----DPNARRCMASAAVAFMTFGDDEBMGYCYDDEHADAVALMGSM 220

Qy 186 DHGYPI-----REHLN-----RLVGDGIWAPAIDGAFVLTTRGDP 222

Db 221 AEMHPIIMTRLRADRLGHEHYKVAVISTFTHRSMDLADIPVKP----- 265
 QY 223 DLQGTDVAI-GY-ASHDTTERL---YLQETLTF 253
 266 ----GTDLAILNTIANHITGTGRVNEDPFAKHITFM 297

RESULT 8

B84331
 hypothetical protein Vng1794c [imported] - Halobacterium sp. NRC-1

C/Species: Halobacterium sp. NRC-1
 C/Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
 C/Accession: B84331
 R/Ref: M.V.; Kennedy, S.P.; Mahatras, G.G.; Bergquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.; Leitner, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jaldic
 Jung, K.H.; Alam, M.; Preltas, T.
 Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
 A/Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li
 A/Title: Genome sequence of Halobacterium species NRC-1.
 A/Reference number: A84160; MUID:20504483; PMID:11016950
 A/Accession: B84331
 A/Status: Preliminary
 A/Molecule type: DNA
 A/Residues: 1-694 <STO>
 A/Cross-references: UNIPROT:Q9HP56; UNIPARC:UPI000006398B; GB:AE004437; NID:G10581248; F
 C/Genetic: A;Gene: VNG1794C

Query Match 7.4%; Score 100; DB 2; Length 694;
 Best Local Similarity 23.1%; Pred. No. 6.2;
 Matches 58; Conservative 38; Mismatches 83; Indels 72; Gaps 14;

QY 11 VTEAAMAEIELEAKRPFKHIAGRVYVS-DEGPPTAAVSGRLIDVKAFTNGVIAHL 69
 346 ITDAALIKARGDAATATARALGGQVTLSPDAGP--ARGDYGRLGYITLPDGTYYNR 403
 QY 70 RA-SKPLVRLVPEFTLSRNEIDD-----VERGSKDSWEPVKAACKLA FVEDRTTFEG 122
 404 RAVADGARY--YASAFSGYDVMABATARAAGDLWE-----LSDPYIDG 449
 Db 123 YSAASIEG-----IRASNPALTLPDPREIPDVISQALSELRLAGVDG 167
 450 PDSRSVSELPFPVAVATGGQTVVASADSTPLVT-----ADAAARVAAVGG 495
 Db 168 PYSVLASADVYTVKSE--TSDHGVPRIEHLNRLVD--GGIIMAPA-IDGAFVLTTRGGD 221
 QY 496 P---LIDGEPFAPSBDGAATNEGHVSMFAVELIDCLADGLSAPVLVDG---HG-- 545
 Db 222 FDLQLGTDVAI 232
 546 ---QPGSDAL 553

RESULT 9

T02553
 cellulose synthase homolog T26B15.10 - Arabidopsis thaliana

C/Species: Arabidopsis thaliana (mouse-ear cress)
 C/Date: 05-Mar-1999 #sequence_revision 05-Mar-1999 #text_change 09-Jul-2004
 C/Accession: T02553; D84734
 R/Rounsaey, S.D.; Kaul, S.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes,
 submitted to the EMBL Data Library, July 1998
 A/Description: Arabidopsis thaliana chromosome II BAC T26B15 genomic sequence.
 A/Reference number: Z14678
 A/Accession: T02553
 A/Status: translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-755 <ROU>

A/Cross-references: UNIPROT:O80891; UNIPARC:UPI00000485B0; EMBL:AC004681; NID:G3298532;
 A/Experimental source: cultivar Columbia
 R/Lin, X.; Kaul, S.; Rounsaey, S.D.; Shee, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
 M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon, L.;
 euse, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.

Nature 402, 761-768, 1999
 A/Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
 A/Reference number: A84420; MUID:20083487; PMID:10617197

A/Accession: D84734
 A/Status: Preliminary
 A/Molecule type: DNA
 A/Residues: 1-755 <STO>
 A/Cross-references: UNIPARC:UPI00000485B0; GB:AE002093; NID:G3298542; PIDN:AA025936.1;
 C/Genetic: A;Gene: AT2G32540; T26B15.10
 A/Map position: 2
 A/Insertions: 88/2; 194/3; 236/3; 277/3; 344/3; 390/1; 450/3; 567/3

Query Match 7.4%; Score 99.5; DB 2; Length 755;
 Best Local Similarity 20.8%; Pred. No. 7.6;
 Matches 71; Conservative 49; Mismatches 112; Indels 109; Gaps 18;

QY 11 VTEAAMAEIELEAKRPFKHIAGR-----RVYVSPD--GGPPTAAVSGRLIDVKA 60
 68 ITYIKSPADY--KTYPERLDERVHELPPVDMFVTADVPREPLIVNTVLSLAVNY 124
 Db 61 PTNGVIAHL-----RASK-----PL-----VRLVPEFTLSRNEIDDERG 95
 125 PANKLACTVSDCCSLTTPSLKASKPKIWPCKKXNVRAAFPMFRNSPELAEGS 184
 QY 96 SKDSWEPVKAACKLA-FVEDRTTFEGYSA--SIEGIRASNSPALTFP----- 144
 185 EFGKQWEMTKREYKLSQKVEDATGSSHWLDAEDFEAFINTKSNDSHTIVKVMENKCG 244
 Db 145 --DPRRIPIVY-----SQALSEL-RLAGV--DGPYVLSADVYTVKSER 184
 245 VGDEKEVPFVYVYSRKRPNHFNHYKAGMNFVRSGLMTNAPYMLNVDCCMYVNEADV 304
 QY 185 -----SDH-----GYP-----IREHLNRLVDGDIIMAPAIDGAF 213
 305 VROAMCTFLQKSDSHKCAVQYPPDPFYSNVGELVQLVYLRGAG--IQSPYRAGSG 362
 Db 214 VLTTRGADPDL---QLGTVGAIGVASHDTTERLYLQETLT 251
 363 CFHTRRVMYGLSLDLDGDDGSLSSIA---TRKYLAESILT 399

RESULT 10

AB2338
 acetyl-coenzyme A synthetase [imported] - Nostoc sp. (strain PCC 7120)

C/Species: Nostoc sp. PCC 7120
 A/Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
 C/Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
 C/Accession: AB2338
 R/Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi,
 Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata,
 DNA Res. 8, 205-213, 2001
 A/Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium An
 A/Reference number: AB1807; MUID:21555285; PMID:11759840
 A/Accession: AB2338
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-524 <RUR>
 A/Cross-references: UNIPROT:Q8YPR0; UNIPARC:UPI000000CEAB1; GB:BA000019; PIDN:BA075956.1
 A/Experimental source: strain PCC 7120
 C/Genetic: A;Gene: al14257

A/Superfamily: acetate-CoA ligase; acetate-CoA ligase homology
 Query Match 7.2%; Score 96.5; DB 2; Length 524;
 Best Local Similarity 25.8%; Pred. No. 7.8;
 Matches 34; Conservative 25; Mismatches 46; Indels 27; Gaps 6;

QY 149 IPDVIQALSELRLAGVDGPVYLSLA-----DYYTVSESTSHGY-----PI 191
 146 MPPIPEAKIATMACAKIAGPHSVVPGFSAELRDLIDAGAVVYVTTADGGMKDAIVPL 205
 Db 192 REHLNRLVDGDIIMAPAIDGAFVLTTRGGDFDLQG-----TDVAIGVASHDTTERLY 245

Db	206	KEQVXKALADDVY--PSVENVLVYKRTGDDITVQLQGRDHMHDLQIG--ASADCPABPMD	262
Qy	246	IQETLTPLCYTA	257
Db	263	-SEDLFPVLYIS	273

RESULT 11

S73015
polyketide synthase pksF - Mycobacterium leprae
N:Alternate names: L518 Pl 8 protein
C:Species: Mycobacterium leprae
C:Date: 19-Mar-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004
C:Accession: S73015
R:Smith, D.R.; Robison, K.
submitted to the EMBL Data Library, November 1993
A:Description: Mycobacterium leprae cosmid L518.
A:Reference number: S72591
A:Accession: S73015
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-1489 <SMI>
A:Cross-references: UNIPROT:Q49934; UNIPARC:UPI00000CA9342; EMBL:U00023; NID:g647194; PIR

RESULT 12

T50734
Regulatory protein ppsr [imported] - Rhodobacter sphaeroides
C:Species: Rhodobacter sphaeroides
C:Date: 21-Jul-2000 #sequence_revision 21-Jul-2000 #text_change 09-Jul-2004
C:Accession: T50734
R:Choudhary, M.; Kaplan, S.
Nucleic Acids Res. 28, 862-867, 2000
A:Title: DNA sequence analysis of the photosynthesis region of Rhodobacter sphaeroides 2
A:Reference number: Z25222; MUID:20115911; PMID:10648776
A:Accession: T50734

A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1464 <CPO>
 A:CROSS-references: UNIPROT:Q53224; UNIPARC:UPI00000B9F09; EMBL:AF195122; PIDD:AAF24278
 A:Experimental source: Strain 2.4.1
 C:Genetics:
 C:Gene: ppsr

RESULT 13

ATP-dependent RNA helicase XF0252 [imported] - Xylella fastidiosa (strain 9a5c)
 C/Species: Xylella fastidiosa
 C/Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
 C/Accession: G82930
 R/anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequencing
 Nature 406, 151-157, 2000
 A/Title: The genome sequence of the plant pathogen Xylella fastidiosa.
 A/Reference number: A82515; MUID:20365717; PMID:10910347
 A/Note: for a complete list of authors see reference number A59328 below
 A/Accession: G82830
 A/Status: preliminary
 A/Molecule type: DNA
 A/Readers: 1-614 <STM>
 A/Cross-references: UNIPROT:Q9F6P6; UNIPARC:UPI00000C234B; GB:AE003878; GB:AE003849; NII
 A/Experimental source: strain 9a5c
 R/Simpson, A.J.G.; Reinisch, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.;
 Britones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carreiro, D.M.; Carreir,
 as-Veto, E.; Docena, C.; El-Dorothy, H.; Facinomi, A.P.; Ferreira, A.J.S.
 submitted to Genbank, June 2000
 A/Authors: Ferreira, V.C.A.; Perro, J.A.; Fraga, J.S.; Franco, S.C.; Franco, M.C.; Froh
 J.D.; Jungueira, M.L.; Kemper, E.L.; Kitzjima, J.P.; Kitzger, J.B.; Kurimae, E.B.; Laig
 Chado, M.A.; Madelira, A.M.B.N.; Madelira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins
 ;A/Authors: Martins, O.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.
 ; F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmetti, D.
 ; Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de sa, R.G.; Santelli, R.V.; Sawasa
 A/Authors: da Silva, A.C.R. de M.; da Silva, P.R. de M.; da Silva, A.S.; Silva Jr., W.A.; da Silve
 M.; Tashiro, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.;
 A/Reference number: A59328
 A/Contents: annotation
 A/Genetics:
 A/Gene: XF0252

Db 89 LAETRELA-----IQVAEAFQRYAASISGRVLPVYGSGSYGOQLALKKGHVHVGTP- 142
 QY 63 NGVIATL-RASKEPLVLRVPEFTLSRNEIDVBERGSKDSDEPVKAACKLAFYEDRTIFE 121
 Db 143 GRVIDHIERKTLDLSEK---TLVLDEADEMLKRGFIED---VEEVLRKLPASRQVAFS 196
 QY 122 GYSAASIEGIRASASNPALTLPEDPREIPVVISQALSELRLAGVGPY---SVLLSADV 178
 Db 197 ATWPPQIRRI-----AQTYYLDPEIR---VTIATKTTTAAAIRQRYYWVWVGILKLDAL 245
 QY 179 TKVISESDHGYPI-----REHNRNVGDDIIMAPALIGAF-----VLTRGDF 222
 Db 246 TRILEETFPDAMIIVRTKATRELAERKIQARGLTAAINQWQOQACRERTIHLKDGKL 305
 QY 223 DLQGTDAIGVASHDTTERL 244
 Db 306 DILVATDV-----AARGIDVERI 323

RESULT 14

A45456
 NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) chain N003 - Paracoccus denitrificans
 C/Species: Paracoccus denitrificans
 C/Date: 24-Feb-1994 #sequence_revision 12-Apr-1996 #text_change 09-Jul-2004
 C/Accession: S23948; A45456
 R/Xu, X.; Matsuno-Yagi, A.; Yagi, T.
 Arch. Biochem. Biophys. 296, 40-48, 1992
 A/Title: Structural features of the 66-kDa subunit of the energy-transducing NADH-ubiquinone
 A/Reference number: S23946; PMID:92296779; PMID:1605643
 A/Accession: S23948
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-673 <XUA>
 A/Cross-references: UNIPROT:P29915; UNIPARC:UPI000016FD3C; GB:M64572; NID:G150599; PIDN:
 R/Xu, X.; Matsuno-Yagi, A.; Yagi, T.
 Biochemistry 32, 968-981, 1993
 A/Title: DNA sequencing of the seven remaining structural genes of the gene cluster encod
 A/Reference number: A45456; PMID:93136200; PMID:8422400
 A/Accession: A45456
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 658-673 <XUI>
 A/Cross-references: UNIPARC:UPI0000175098
 A/Note: sequence extracted from NCBI backbone (NCBI:123409, NCBI:123410)
 C/Superfamily: NADH dehydrogenase (ubiquinone) chain 11
 C/Keywords: NAD; oxidoreductase

Query Match 6.8%; Score 93.5; DB 2; Length 673;
 Best Local Similarity 22.8%; Pred. No. 19;
 Matches 66; Conservative 45; Mismatches 119; Indels 59; Gaps 14;

QY 6 RDLAPVTEAMARELEAARTPKRHINGRVVSDPGPVTAAVSTGRLLID-----VKA 60
 Db 290 RENGRLKPAPWPELEPAAPAMK---GKTIAGLIGDLPAPAAAFSLKQVTEGLGKVEEC 345
 QY 61 PTNGVIALRLASKPLVLRVPEFTLSRNEIDVBERGSKDSDEPVKAACKLAFYED 116
 Db 346 RVNG--ARLPAGN---RSAYVGTARIEDIDDAEMIQIGTNPRDEAVLNARIRKASWG 400
 QY 117 RTI-----FEGYSAASIEGIRASASNPALTLPEDPREIPVVISQALSELRL 162
 Db 401 AKVGLVGEVPDLYFDYVAVHGTDRALRLSSRR-----ISDETKARPSIV--IVCGAI 452
 QY 163 AGVDGPFYSVLSADVYTKVSETSDHGYPIRE-----HNRNVGDDIIMAPALIGAF 213
 Db 453 ARBDGE-AVLAHA---MKLENSNSGLLIHTTAAGRVGANDVGAIVTEGGL--AAIDGAB 506
 QY 214 VLTRGSD-FDLQGTDAIGVASHDTTERLYLQETLTPFLCYTAASV 261
 Db 507 VVNLGADVEDYDIOGPFVIVOGSHGDRGAHRDI---LPGACTEBSGL 552

RESULT 15

B70982
 probable nei protein - Mycobacterium tuberculosis (strain H37Rv)
 C/Species: Mycobacterium tuberculosis
 C/Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
 C/Accession: B70982
 R/Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.;
 Rajandream, M.A.; Davies, R.; Devlin, K.; Feltham, T.; Gentles, S.; Hamlin, N.; Holroyd, S.
 Nature 393, 537-544, 1998
 A/Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
 A/Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
 A/Reference number: A70500; PMID:98295987; PMID:9634230
 A/Accession: B70982
 A/Status: preliminary; nucleic acid sequence not shown; translation not shown
 A/Molecule type: DNA
 A/Residues: 1-255 <COL>
 A/Cross-references: UNIPROT:P96902; UNIPARC:UPI0000129EDB; GB:Z92771; GB:AL123456; NID:
 A/Experimental source: strain H37Rv
 C/Genetics:
 A/Gene: nei
 C/Superfamily: formamidyrimidine-DNA glycosylase

Query Match 6.8%; Score 92; DB 2; Length 255;
 Best Local Similarity 21.8%; Pred. No. 6.5;
 Matches 62; Conservative 39; Mismatches 94; Indels 90; Gaps 14;
 QY 10 PTEAMARELEAARTPKRHINGRVV--DVSDP-----GGPVTAAVSTGRLLIDVK 59
 Db 2 PBDYTW-----HTAATLRHLAAGRLTTCIDIRVPRFAVVDLGEVVDVLSGKHLFIR 56
 QY 60 APTNGVIALRL-----ASKPL-VLRVPEFTLSRNE-----IDVBERGSK 97
 Db 57 TGTASIHSLQMDGSRVGNRPVPRVDHRRARILLEANQOEQARVGVVDLGLVIRHND 116
 QY 98 DS-----DMEPVKAACKLAFYEDRTIFEGYSAASIEGIRASASNPALTLPED 145
 Db 117 GAVALHLPDLADMDP--QRAAANLIVAPDRILAE-----ALL---D 155
 QY 146 PREIPVISQALSEL-RLAGVDGPFYSVLSADVYTKVSETSDHGYPIREHNLVGDGI- 203
 Db 156 QRVLAGIGNVYCNELCFVSGVLPFTAPVSAVADPRRLVTRABDMLVNRPRMNRCTTGDFR 215
 QY 204 ----IAPALDGAFTVLTTRGSDPDLQGTDAIGVASHDTTERL 244
 Db 216 AGRRLW-----VYGRAGGCCRRCGTLIA-----YDTDERV 246

Search completed: April 14, 2006, 17:34:23
 Job time : 18.539 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: April 14, 2006, 17:04:10 ; Search time 97.7763 Seconds
(without alignments)
1912.171 Million cell updates/sec

Title: US-10-620-246-16

Perfect score: 1348
Sequence: 1 MNNLYRDLAPVTBAABAEIR.....LQETTLFPLCYTAASVALSH 265

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_05.80.*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1348	100.0	265	007812_MYCTU	007812 mycobacteri
2	1341	99.5	265	07U190_MYCBO	07U190 mycobacteri
3	1341	99.5	265	007181_MYCTU	007181 mycobacteri
4	1188	88.1	265	074385_MYCPA	074385 mycobacteri
5	914.5	67.8	266	05YPI3_NOCRA	05YPI3 nocardia fa
6	840	62.3	267	07WAG7_BORBP	07WAG7 bordetella
7	840	62.3	267	07WJ14_BORBR	07WJ14 bordetella
8	823	61.1	271	04LRT9_BORBR	04LRT9 burkholderi
9	780.5	57.9	266	045296_BRBLN	045296 brevibacter
10	364	28.5	262	09WZP2_THEMA	09WZP2 thermotoga
11	365.5	27.1	266	07MSH9_WOLSU	07MSH9 wolliella s
12	196.5	14.6	284	05L1H9_GEOXA	05L1H9 geobacillus
13	187.5	13.9	272	04NSX8_9DELT	04NSX8 anaeromyxob
14	142	10.5	281	067639_AQUAE	067639 aquilex aeo
15	112	8.3	453	026970_TRYCR	026970 trypanosoma
16	112	8.3	492	1_TYTR_TRYCR	228553 trypanosoma
17	108.5	8.0	917	06N4Z6_RHOPA	06N4Z6 rhodopseudo
18	106.5	7.9	1293	07QSN9_GIALA	07QSN9 giardia lam
19	106.5	7.9	1293	08MUV2_GIALA	08MUV2 giardia lam
20	105.5	7.8	390	09YEN2_AERPE	09YEN2 aeropyrum p
21	105.5	7.8	467	1_Y1435_DELRA	09YEN2 aeropyrum p
22	104.5	7.8	508	055YA3_CRYNE	055YA3 cryprococcu
23	104.5	7.8	508	05KLM5_CRYNE	05KLM5 cryprococcu
24	103	7.6	805	07KR73_DROME	07KR73 drosophilla
25	102.5	7.6	834	0922Z6_RHIME	0922Z6 rhizobium m
26	101.5	7.5	873	07ZAH0_3CRAN	07ZAH0 acidithiobac
27	101	7.5	554	2_09KJAO_ECOLI	09KJAO escherichia
28	101	7.5	554	2_0842D0_BCOI1	0842D0 escherichia
29	101	7.5	696	2_08MS23_DROME	08MS23 drosophilla
30	100	7.4	694	2_09HP56_HALSA	09HP56 halobacteri
31	99.5	7.4	422	2_08LIX4_ORYSA	08LIX4 oryza sativ

32	99.5	7.4	755	2	080891_ARATH	080891 arabidopsis
33	99	7.3	588	2	05UXK7_HALMA	05UXK7 halocaula
34	99	7.3	762	2	07VTQ8_BORBP	07VTQ8 bordetella
35	99	7.3	762	2	07VZ68_BORBP	07VZ68 bordetella
36	98.5	7.3	271	2	092SL4_RHIME	092SL4 rhizobium m
37	98.5	7.3	456	2	07UK29_RHOBA	07UK29 rhodospirill
38	98	7.3	394	2	04KN75_BORBR	04KN75 bordetella
39	98	7.3	410	2	0827X3_STRAW	0827X3 streptomyc
40	98	7.3	762	2	07WNV9_BORBR	07WNV9 bordetella
41	97.5	7.2	2221	2	087P88_VIBPA	087P88 vibrio para
42	97	7.2	403	2	04TOR6_9SPBN	04TOR6 erythrobact
43	97	7.2	531	2	09BE00_RHITO	09BE00 rhizobium l
44	97	7.2	766	2	08H5B0_MAIZE	08H5B0 zea mays (m
45	97	7.2	786	2	06DW74_LOTUA	06DW74 lotus japon

ALIGNMENTS

RESULT 1

007812_MYCTU PRELIMINARY; PRT; 265 AA.

AC 007812_01-JUN-1997 (TREMBLrel. 04, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE 29 Kda antigen.
GN Name=cfp29;
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacteriaceae; Mycobacteriaceae; Mycobacterium;
OC Mycobacterium tuberculosis complex.
OX NCBI_TaxID=1773;
RN [1]
RP NCBIOTIDE SEQUENCE.
RC STRAIN=H37Rv;
RX MDLINE=98261519; PubMed=9596740;
RA Rosenkranz I., Rasmussen P.B., Carnio M., Jacobsen S., Theisen M.,
RA Andersen P.;
RT "Identification and characterization of a 29-kilodalton protein from
RT Mycobacterium tuberculosis culture filtrate recognized by mouse memory
RT effector cells.";
RL Infect. Immun. 66:2728-2735 (1998).
DR EMBL; Y12820; CAA73350.1; -; Genomic DNA.
DR InterPro; IPR007544; Linocin_M18.
DR Pfam; PR0454; Linocin_M18; I.
SQ SEQUENCE 265 AA; 28860 MW; A1C4B96904D3B325 CRC64;

Query Match 100.0%; Score 1348; DB 2; Length 265;
Best Local Similarity 100.0%; Pred. No. 1.2e-94;
Matches 265; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MNNLYRDLAPVTBAABAEIRLRAATPKRIAGRRVVDSDPGPPTAAVSTGRLLIDVKA	60
DB	1	MNNLYRDLAPVTBAABAEIRLRAATPKRIAGRRVVDSDPGPPTAAVSTGRLLIDVKA	60
QY	61	PTNGVIAHLPASKEVLRIRVPTLSRNEIDVRSKSDSMEVVKAAKGLAVERTIP	120
DB	61	PTNGVIAHLPASKEVLRIRVPTLSRNEIDVRSKSDSMEVVKAAKGLAVERTIP	120
QY	121	EGSASASIEGIRASNPALTLPEDPREIPDVISOALSELRLAGVGPVSLSDVYTK	180
DB	121	EGSASASIEGIRASNPALTLPEDPREIPDVISOALSELRLAGVGPVSLSDVYTK	180
QY	181	VSETSHGVPRIHRLNRLVDGDIIPAPIDAFAVLTTRGGDFLQGTDAIYASHDT	240
DB	181	VSETSHGVPRIHRLNRLVDGDIIPAPIDAFAVLTTRGGDFLQGTDAIYASHDT	240
QY	241	TERLYVQETTLFPLCYTAASVALSH	265
DB	241	TERLYVQETTLFPLCYTAASVALSH	265

RESULT 2
 Q70190 MYCBO PRELIMINARY; PRT; 265 AA.
 ID Q70190 MYCBO PRELIMINARY; PRT; 265 AA.
 AC Q70190
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE 29 kDa ANTIGEN CFP29.
 GN Name=cfp29; OrderedLocuNames=MB0821C;
 OS Mycobacterium bovis.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacteriaceae; Mycobacteriaceae; Mycobacterium;
 OC Mycobacterium tuberculosis complex.
 OX NCBI_TaxID=1765;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=AF2122/97;
 RX MEDLINE=22709107; PubMed=12788972; DOI=10.1073/pnas.1130426100;
 RA Garnier T., Eglmeier K., Camus J.-C., Medina N., Mansoor H.,
 RA Payor M., Duthey S., Grondin S., Lacroix C., Monsemp C., Simon S.,
 RA Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,
 RA Parthill J., Barrell B.G., Cole S.T., Gordon S.V., Hewinson R.G.,
 RT "The complete genome sequence of Mycobacterium bovis."
 RL Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882(2003).
 DR EMBL; BX248336; CAD93683.1; -; Genomic_DNA.
 DR InterPro; IPR007544; Lincocin_M18.
 DR Pfam; PF04454; Lincocin_M18; 1.
 DR Complete proteome.
 SQ SEQUENCE 265 AA; 28830 MW; 00E7E96904D3B33A CRC64;

Query Match 99.5%; Score 1341; DB 2; Length 265;
 Best Local Similarity 99.6%; Pred. No. 4e-94;
 Matches 264; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MNLYRDLAVPTAAAEIELEAARTFKRHIAAGRVVDSDPGPVTAAVSTRLLIDYKA 60
 DB 1 MNLYRDLAVPTAAAEIELEAARTFKRHIAAGRVVDSDPGPVTAAVSTRLLIDYKA 60
 QY 61 PTNGVIAHLRASPLVRLVPTLSRNEIDVERGSKDSQDWEPEKAAKKAFAVEDRTIF 120
 DB 61 PTNGVIAHLRASPLVRLVPTLSRNEIDVERGSKDSQDWEPEKAAKKAFAVEDRTIF 120
 QY 121 EGYSAASIEGIRASSNPALTLPEDPREIPDVISOALSRLAGVDPYSVLLSADVYTK 180
 DB 121 EGYSAASIEGIRASSNPALTLPEDPREIPDVISOALSRLAGVDPYSVLLSADVYTK 180
 QY 181 VSETSDHGYPYRHLRLVVDGDIIMAPALDGAFLVLTTRGGDFDLQGTGVAIGYASHDT 240
 DB 181 VSETSDHGYPYRHLRLVVDGDIIMAPALDGAFLVLTTRGGDFDLQGTGVAIGYASHDT 240
 QY 241 TERLYLOETLTFPCYTAASVALSH 265
 DB 241 TERLYLOETLTFPCYTAASVALSH 265

RESULT 3
 Q07181 MYCTU PRELIMINARY; PRT; 265 AA.
 ID Q07181 MYCTU PRELIMINARY; PRT; 265 AA.
 AC Q07181; Q7D997;
 DT 01-JUL-1997 (TrEMBLrel. 04, Created)
 DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
 DE 29 kDa ANTIGEN CFP29 (Bacteriocin CFP29).
 GN Name=cfp29; OrderedLocuNames=MT0819, Rv0798C;
 OS Mycobacterium tuberculosis.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacteriaceae; Mycobacteriaceae; Mycobacterium;
 OC Mycobacterium tuberculosis complex.
 OX NCBI_TaxID=1773;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=H37RV;
 RX MEDLINE=98295987; PubMed=9634230; DOI=10.1038/31159;

RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C.M.,
 RA Harris D.E., Gordon S.V., Eigmeier K., Gas S., Barry C.E. III,
 RA Tekala F., Badcock K., Basham D., Brown D., Chillingworth T.,
 RA Connor R., Davies R.M., Devlin K., Feldwell T., Gentles S., Hamlin N.,
 RA Holroyd S., Hornsby T., Jagels K., Krogh A., McLean J., Moule S.,
 RA Murphy L.D., Oliver S., Osborne J., Quail M.A., Rajandream M.A.,
 RA Rogers J., Rutter S., Seeger K., Skellern S., Squares R.,
 RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.,
 RT "Deciphering the biology of Mycobacterium tuberculosis from the
 RT complete genome sequence."
 RL Nature 393:537-544(1998).
 RN [2]

RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=CDC 1551 / Oshkosh;
 RX MEDLINE=2206494; PubMed=12218036;
 RX DOI=10.1128/JB.184.19.5479-5490.2002;
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
 RA Peterson J.D., DeBoy R.T., Dodson R.J., Gait M.L., Haft D.H.,
 RA Hickey E.K., Kolonay J.F., Nelson W.C., Umayam L.A., Esmolaeva M.D.,
 RA Salzberg S.L., Delcher A., Utecherback T.R., Weidman J.F., Kouri H.M.,
 RA Gill J., Mikula A., Bishai W., Jacobs W.R. Jr., Venter J.C.,
 RA Fraser C.M.;
 RT "Whole-genome comparison of Mycobacterium tuberculosis clinical and
 RT laboratory strains."
 RL J. Bacteriol. 184:5479-5490(2002).
 DR EMBL; BX42574; CAB09572.1; -; Genomic_DNA.
 DR EMBL; AB000516; AAK45061.1; -; Genomic_DNA.
 DR PIR; E70520; E70520.
 DR TIGR; MT0819; -;
 DR TubercuList; Rv0798C; -;
 DR InterPro; IPR007544; Lincocin_M18.
 DR Pfam; PF04454; Lincocin_M18; 1.
 DR Complete proteome.
 SQ SEQUENCE 265 AA; 28830 MW; 00E7E96904D3B33A CRC64;

Query Match 99.5%; Score 1341; DB 2; Length 265;
 Best Local Similarity 99.6%; Pred. No. 4e-94;
 Matches 264; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MNLYRDLAVPTAAAEIELEAARTFKRHIAAGRVVDSDPGPVTAAVSTRLLIDYKA 60
 DB 1 MNLYRDLAVPTAAAEIELEAARTFKRHIAAGRVVDSDPGPVTAAVSTRLLIDYKA 60
 QY 61 PTNGVIAHLRASPLVRLVPTLSRNEIDVERGSKDSQDWEPEKAAKKAFAVEDRTIF 120
 DB 61 PTNGVIAHLRASPLVRLVPTLSRNEIDVERGSKDSQDWEPEKAAKKAFAVEDRTIF 120
 QY 121 EGYSAASIEGIRASSNPALTLPEDPREIPDVISOALSRLAGVDPYSVLLSADVYTK 180
 DB 121 EGYSAASIEGIRASSNPALTLPEDPREIPDVISOALSRLAGVDPYSVLLSADVYTK 180
 QY 181 VSETSDHGYPYRHLRLVVDGDIIMAPALDGAFLVLTTRGGDFDLQGTGVAIGYASHDT 240
 DB 181 VSETSDHGYPYRHLRLVVDGDIIMAPALDGAFLVLTTRGGDFDLQGTGVAIGYASHDT 240
 QY 241 TERLYLOETLTFPCYTAASVALSH 265
 DB 241 TERLYLOETLTFPCYTAASVALSH 265

RESULT 4
 Q743F5 MYCPA PRELIMINARY; PRT; 265 AA.
 ID Q743F5 MYCPA PRELIMINARY; PRT; 265 AA.
 AC Q743F5;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Hypothetical protein.
 GN OrderedLocuNames=MA0630C;
 OS Mycobacterium paratuberculosis.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacteriaceae; Mycobacteriaceae; Mycobacterium;
 OC Mycobacterium avium complex (MAC).


```

OX NCBI_TaxID=1770;
RN (1)
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=K10;
RA Li L., Bannantine J., Zhang Q., Amonsin A., Alt D., Kapur V.;
RL Submitted (SRP-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL: AEO17229; AAS02947.1; -; Genomic_DNA.
DR InterPro: IPR007544; Lincocin_M18.
DR Pfam: PF04454; Lincocin_M18; 1.
KM Complete proteome.
SQ SEQUENCE 265 AA; 28618 MW; 6157F708A7DC98C9 CRC64;

Query Match 88.1%; Score 1188; DB 2; Length 265;
Best Local Similarity 86.4%; Pred. No. 1.9e-82;
Matches 228; Conservative 20; Mismatches 16; Indels 0; Gaps 0;

QY 1 MNNUYDLAPVTEAAMAEIELEAARTFKRIHAGRVVDVSDPGPVTAAVSTGLIDVKA 60
DB 1 MNNUYDLAPVTEAAMAEIELEAARTFKRIHAGRVVDVSDPGPAAAVSTGLIDVEA 60
QY PTNGVIAHLASKPLVLRVPFTLSRNEIDVERGSKDSMPVKEAAKTLAFVEDRTIF 120
DB PTNGVIAHLASKPLVLRVPFTLSRNEIDVERGANDSDMPVKEAAKTLAFVEDRAIF 120
QY 61 PTNGVIAHLASKPLVLRVPFTLSRNEIDVERGSKDSMPVKEAAKTLAFVEDRTIF 120
DB 61 PTNGVIAHLASKPLVLRVPFTLSRNEIDVERGANDSDMPVKEAAKTLAFVEDRAIF 120
QY 121 EGYSAASISIGIRASNSPALTLPEDPREIPDVISOALSRLAGVDGSPYVLSADVYTK 180
DB 121 EGYSAASISIGIRASNSKPLALPADPRIPDVITQALISRLAGVDGSPVLSADVYTK 180
QY 121 EGYSAASISIGIRASNSPALTLPEDPREIPDVISOALSRLAGVDGSPYVLSADVYTK 180
DB 121 EGYSAASISIGIRASNSKPLALPADPRIPDVITQALISRLAGVDGSPVLSADVYTK 180
QY 181 VSETSDHGYPRIHRLNRLV-DGDIWAPAIDGAFVLTTRGGDFDLOGTDAVIGYASHDT 240
DB 181 VSETSDHGYPRIHRLNRLV-DGDIWAPAIDGAFVLTTRGGDFDLOGTDAVIGYASHDT 240
QY 181 VSETSDHGYPRIHRLNRLV-DGDIWAPAIDGAFVLTTRGGDFDLOGTDAVIGYASHDT 240
DB 181 VSETSDHGYPRIHRLNRLV-DGDIWAPAIDGAFVLTTRGGDFDLOGTDAVIGYASHDT 240
QY 241 TERLYLQETLTFPCYTAASVALS 264
DB 241 TVQLYLQETLTFPCYTAASVPLT 264

RESULT 5
Q5YPL3 NOCPA
ID Q5YPL3_NOCPA PRELIMINARY; PRT; 266 AA.
AC Q5YPL3;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
GN Putative bacteriocin family protein.
DN OrderedLocustNames=efsa50260;
OS Nocardia farcinica.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacteriaceae; Nocardiaceae; Nocardia.
OX NCBI_TaxID=37329;
RN (1)
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=IFM 10152;
RX PubMed=15466710; DOI=10.1073/pnas.0406410101;
RA Ishikawa J., Yamaehita A., Mikami Y., Hoshino Y., Kurita H., Hotta K.,
RA Shiba T., Hattori M.;
RL "The complete genomic sequence of Nocardia farcinica IFM 10152."
RL Proc. Natl. Acad. Sci. U.S.A. 101:14925-14930(2004).
DR EMBL: APO06618; BAD59878.1; -; Genomic_DNA.
DR InterPro: IPR007544; Lincocin_M18.
DR Pfam: PF04454; Lincocin_M18; 1.
KM Complete proteome.
SQ SEQUENCE 266 AA; 28763 MW; D77C77B24P4F73FA CRC64;

Query Match 67.8%; Score 914.5; DB 2; Length 266;
Best Local Similarity 67.4%; Pred. No. 1.4e-61;
Matches 178; Conservative 35; Mismatches 50; Indels 1; Gaps 1;

QY 1 MNNUYDLAPVTEAAMAEIELEAARTFKRIHAGRVVDVSDPGPVTAAVSTGLIDVKA 60
DB 1 MNNUYDLAPVTEAAMAEIELEAARTFKRIHAGRVVDVSDPGPVTAAVSTGLIDVKA 60
QY 1 MNNUYDLAPVTEAAMAEIELEAARTFKRIHAGRVVDVSDPGPVTAAVSTGLIDVKA 60
DB 1 MNNUYDLAPVTEAAMAEIELEAARTFKRIHAGRVVDVSDPGPVTAAVSTGLIDVKA 60
QY 61 PTNGVIAHLASKPLVLRVPFTLSRNEIDVERGSKDSMPVKEAAKTLAFVEDRTIF 120
```

```

DB 61 PDEGVQARQVRVAPLVLELRVPFTLSRNEIDVERGANDTDLDAVKEAARRIAFEDRAIF 120
QY 121 EGYSAASISIGIRASNSPALTLPEDPREIPDVISOALSRLAGVDGSPYVLSADVYTK 180
DB 121 EGYPAAGITGIRASNSAPLTTPVEDDRLVREBALTOALRALAGVDGSPYVLSABEYTE 180
QY 181 VSETSDHGYPRIHRLNRLV-DGDIWAPAIDGAFVLTTRGGDFDLOGTDAVIGYASHDT 239
DB 181 VSETSDHGYPRIHRLNRLV-DGDIWAPAIDGAFVLTTRGGDFDLOGTDAVIGYASHDT 240
QY 240 DTERLYLQETLTFPCYTAASVAL 263
DB 241 DTERLYLQETLTFPCYTAASVAL 264

RESULT 6
Q7WAG7 BORPA
ID Q7WAG7 BORPA PRELIMINARY; PRT; 267 AA.
AC Q7WAG7;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DN Bacteriocin.
GN Name=LincM18; OrderedLocustNames=BP1410;
OS Bordetella parapertussis.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Bordetella.
OX NCBI_TaxID=519;
RN (1)
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=12822 / ATCC BAA-587;
RX MEDLINE=22827954; PubMed=12910271; DOI=10.1038/ng1227;
RA Parkhill J., Sebatina M., Preston A., Murphy L.D., Thomson N.R.,
RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
RA Cerdano-Tarraga A.-M., Temple L., James K.D., Hattley B., Quail M.A.,
RA Achtman M., Atkin R., Baker S., Baeham D., Bason N., Cherevach I.,
RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RA Felwell T., Goble A., Hamlin N., Hauser H., Holtroyd S., Jagers K.,
RA Leather S., Moule S., Norberczak H., O'Neill S., Ormond D., Price C.,
RA Rabinowitch B., Rutter S., Sanders M., Saunders D., Seeger K.,
RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
RA Unwin L., Whitehead S., Barrett B.G., Maskell D.J.;
RL "Comparative analysis of the genome sequences of Bordetella pertussis,
RL Bordetella parapertussis and Bordetella bronchiseptica."
RL Nat. Genet. 35:32-40(2003).
DR EMBL: BX640427; CAB36712.1; -; Genomic_DNA.
DR InterPro: IPR007544; Lincocin_M18.
DR Pfam: PF04454; Lincocin_M18; 1.
KM Complete proteome.
SQ SEQUENCE 267 AA; 28506 MW; BEB667BD81DBE25 CRC64;

Query Match 62.3%; Score 840; DB 2; Length 267;
Best Local Similarity 62.5%; Pred. No. 6.9e-56;
Matches 165; Conservative 31; Mismatches 68; Indels 0; Gaps 0;

QY 1 MNNUYDLAPVTEAAMAEIELEAARTFKRIHAGRVVDVSDPGPVTAAVSTGLIDVKA 60
DB 1 MNNUYDLAPVTEAAMAEIELEAARTFKRIHAGRVVDVSDPGPVTAAVSTGLIDVKA 60
QY 1 MNNUYDLAPVTEAAMAEIELEAARTFKRIHAGRVVDVSDPGPVTAAVSTGLIDVKA 60
DB 1 MNNUYDLAPVTEAAMAEIELEAARTFKRIHAGRVVDVSDPGPVTAAVSTGLIDVKA 60
QY 61 PTNGVIAHLASKPLVLRVPFTLSRNEIDVERGSKDSMPVKEAAKTLAFVEDRTIF 120
DB 61 PTNGVIAHLASKPLVLRVPFTLSRNEIDVERGSKDSMPVKEAAKTLAFVEDRTIF 120
QY 61 PTNGVIAHLASKPLVLRVPFTLSRNEIDVERGSKDSMPVKEAAKTLAFVEDRTIF 120
DB 61 PTNGVIAHLASKPLVLRVPFTLSRNEIDVERGSKDSMPVKEAAKTLAFVEDRTIF 120
QY 121 EGYSAASISIGIRASNSPALTLPEDPREIPDVISOALSRLAGVDGSPYVLSADVYTK 180
DB 121 EGYPAAGITGIRASNSAPLTTPVEDDRLVREBALTOALRALAGVDGSPYVLSABEYTE 180
QY 181 VSETSDHGYPRIHRLNRLV-DGDIWAPAIDGAFVLTTRGGDFDLOGTDAVIGYASHDT 240
DB 181 VSETSDHGYPRIHRLNRLV-DGDIWAPAIDGAFVLTTRGGDFDLOGTDAVIGYASHDT 240
QY 241 TERLYLQETLTFPCYTAASVALS 264
```


RL Appl. Environ. Microbiol. 62:1283-1286(1996).
 DR EMBL: X3588; CAA63787.1; -: Genomic_DNA.
 DR InterPro: IPR007544; Litocin_M18.
 DR InterPro: IPR001059; Pept_cys_AS.
 DR Pfam: PF04454; Litocin_M18; 1.
 DR PROSITE: PS00639; THIOL_PROTEASE_HIS; UNKNOWN_1.
 SQ PROSITE: 266 AA; 28597 MM; C33960AC4C05295 CRG64;
 SEQUENCE

Query Match	57.9%	Score 780.5;	DB 2;	Length 266;
Best Local Similarity	58.1%;	Pred. No. 2.4e-51;		
Matches 154;	Conservative 43;	Mismatches 67;	Indels 1;	Gaps 1;

```

QY 1 NNNLFRDLAPVTEBAAMAEIELEAKAFPEKHHINGRRVVVSDPCGPGTAAVSTGRLLIDYKA 60
DB 1 NNNLFRDLAPVTEBAAMAEIELEAKAFPEKHHINGRRVVVSDPCGPGTAAVSTGRLLIDYKA 60
QY 61 PTNGVIAHLRASKPLVRLEVPPTLSHNEIDVERGSKDSDBEYKEAKKLAFEVEDRTIE 120
DB 61 ETSGLQVQKRIQVQVEYIELRTPEFTVTRQALIDVAVRGSGSDWQPCXDAATTTAAEDRALTI 120
QY 121 EYSAASIEGIGISASNSPALITLPEDBREILPDVISOALSLELRLAGVQPGVSYLSADNVYRK 180
DB 121 HGLDAAIGGGIYIPGSSNAVAVALPDVAEDPADADVAQALSVYLRIVGVQPGVSYLSASSAYYRK 180
QY 181 VSESTDHGYPPIREHILNR-LVDDGDIWAIPAIDGAFVLTTRGCGPBDLQIGTDVAIGVASHDT 239
DB 181 VSESTDHGYPPIREHILNR-LVDDGDIWAIPAIDGAFVLTTRGCGPBDLQIGTDVAIGVASHDT 239
QY 240 DTERLYIQETLTFLCTTAASVALS 264
DB 241 ETVELYIQETFGFLALTDRESSVPLS 265

```

RESULT 10	
Q9WZP2_THEME	
ID	Q9WZP2_THEME PRELIMINARY; PRT; 262 AA.
AC	Q9WZP2;
DT	01-NOV-1999 (TEMBIrel. 12, Created)
DT	01-NOV-1999 (TEMBIrel. 12, Last sequence update)
DT	01-JUN-2003 (TEMBIrel. 24, Last annotation update)
DE	Bacteriocin.
GN	OrderedLocusNames=TM0785;
OS	Thermotoga maritima.
OC	Bacteria; Thermotogae; Thermotogales; Thermotogaceae; Thermotoga
OX	NCBI_taxonomyID=2336;

RP NUCLEOTIDE SEQUENCE.
RC STRAIN:MS98 / DSM 3109 / ATCC 43589;
RX MEDLINE=99287316; PubMed=10360571; DOI=10.1038/20601;
RA Nelson K.B., Clayton R.A., Gill S.R., Gwin M.L., Dodson R.J.,
RA Hat D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
RA McDonald L.A., Utechtack T.R., Melek J.A., Linher K.D., Garrett M.M.,
RA Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.L.,
RA Heidelberg J.F., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
RA Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;
RT "Evidence for lateral gene transfer between Archaea and Bacteria from
RT genome sequence of *Thermococcus maritima*.";
RL Nature 399:323-329 (1999).
DR EMBL; AB001747; AACD5867.1; -; Genomic_DNA.
DR PIR; B72333; B72333.
DR TIGR; TM0785; -;
DR InterPro:IPR00544; Lncocin_M18.
DR Pfam; PF04454; lncocin_M18; 1.
KM Complete proteome.
SQ Sequence 262 AA; 30163 MW; 9D45F9B076A97A12 CRC64;

Query Match	28.5%;	Score 384;	DB 2;	Length 262;
Best Local Similarity	36.2%;	Pred. No. 4.2e-21;		
Matches 92;	Conservative 48;	Mismatches 108;	Indels 6;	Gaps 6;

Qy 1 MNNLYRDLAPVTEAAMAEIETEAARTPKRHIAGRVVDSDPGGPVTAAVSTGRLIDVKA 60
| | | | | : | | | | | : | | : | :
Db 4 MEFLLKSPAPLTBKQMGIDNRARBEIFKTQLYGRKFDVDEGPGYGYAAPHLG-VEVLS 62

```

QY 61 PINGIA-HLRASKLVLRVFTLSRNEIDVEGSKSDMEPVBAKKLAFVDRTI 119
DB 63 DENEVWKKGLRSILPLIELRATFTLDLWELDNLERKPNVDLSLEIVAKVAFDEVI 122
QY 120 FEGYSAASIEGIRASNSPALTFPEDPREIPVISOALSRLAGVGPYVLLSDVYT 179
DB 123 FRGCKSGVKKGLSTSEER-KIECGSTPMOLALVRLALSIFSMGIEGGYTLVINDRWI 181
QY 180 K-VSETSDHGYPRIEHLNR-LVDSDIIMAPADGAFVLTTRGSDPQLGSTDVAIGYASH 237
DB 182 NFLKEAGH-YELERKVERECLRGKRIITPRIEDALVSERGDPFLILIGODLSIGYEDR 240
QY 238 DTDTERLYLOETLT 251
DB 241 EKDVAFLFTYETFT 254

```

Q7MSM9_WOLSTU	Q7MSM9_WOLSTU PRELIMINARY,	PRT;	266 AA.
ID	Q7MSM9?		
AC	Q7MSM9?		
DT	01-MAR-2004 (TrEMBLrel. 26, Created)		
DT	01-MAR-2004 (TrEMBLrel. 26, last sequence update)		
DT	01-MAR-2004 (TrEMBLrel. 26, last annotation update)		
DE	LINOCIN M18.		
GN	OrderedlocusNames=MS0287;		
OS	Molinitella succinogenes.		
OC	Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;		
OC	Helicobacteraceae; Molinitella.		
OX	NCBI_TaxID=844;		

RP NUCLEOSIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=DSMZ 1740;
 RX MEDLINE=228828937; PubMed=14500908; DOI=10.1073/pnas.1992838100;
 RA Baar C., Eppinger M., Raddeatz G., Simon J., Lanz C., Klimek O.,
 RA Nandakumar R., Grosse R., Rosinas A., Keller H., Jagtap P., Lähke B.,
 RA Meyer F., Lederer H., Schuster S.C.;
 RT "Complete genome sequence and analysis of *Mollicella succinogenes*.";
 RL Proc. Natl. Acad. Sci. U.S.A. 100(11):1690-11695(2003).
 DR EMBL; BX571657; CA609438.1; -; Genomic_DNA.
 DR InterPro; IPR007544; Interprocln M18.
 DR Pfam; PF04454; linocln_M18; I.
 KW Complete proteome.
 SQ SEQUENCE 266 AA; 28856 MW; 25BDBFB804FEC806 CRC64;

Query Match	27.1%;	Score 365.5;	DB 2;	Length 266;
Best Local Similarity	31.6%;	Pred. No. 1.1e-19;		
Matches	84;	Conservative	60;	Mismatches 117;
			Indels	5;
			Gaps	4

QY	1	NNNNYRDLAPTEAMMAEIEEAATPEKHJLGRVVYVDS	DGGPPTAAVSTGRLLDYKA	60
		1		
Db	1	MDILRENAQPPASWASIEKAGLVFGHJLGRKKVDDPKGLGIGFSSLPTRIVSSKE		60
		1		
QY	61	PTNGVIATLRASKPLVRLRVPEFTLSRMEIDIVERGSKDSWDEPVYKAAKKAFAVEDRTTF		120
		61		
Db	61	KLGBASVGVRRMNTPVYELKIPSPSESEVEATLREANAPDISIEKAAKKVCAENELVF		120
		61		
QY	121	EGYSAASIEGIRKSASSNAPLTPEDBPREFDVYISQALSTRLAGVDPFYSVADVYTK		180
		121		
Db	121	YGLKKKEGLE--IPSIHPKPIKAKGDELPAVAVGRIKELVNSEKGPYALLIQOYFGK		178
		121		
QY	181	V-SETSDHGYPRIEHLNRLVNDG-DYIMAPAI-DGAFVLTTRGSDPDLQLOTDAVAGYASH		237
		181		
Db	179	LEGVAGNSGGYPLTLKLAEELQGNNTIVAPALSKALLVSLRGDGYELVSGMDIGVYSEK		238
		179		
QY	238	DTDTBERLYLOETLTPLCYTAASVAL		263
		238		
Db	239	KSTNHLEPFEEYTLFRINTPEKASIAI		264
		239		

RESULT 12
Q5L1H9_GEOKA

ID 05LJH9_GEOKA PRELIMINARY; PRT; 284 AA.
 AC 05LJH9;
 DT 01-FEB-2005 (TrEMBLrel. 29, Created)
 DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
 DE Hypothetical conserved protein.
 GN OrderedLocuNames=GK0916;
 OS Geobacillus kaustophilus.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Geobacillus.
 NCBI_TaxID=1462;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=HTA426;
 RX PubMed=15576355; DOI=10.1093/nar/gkh970;
 RA Takami H., Takaki Y., Chee G.-J., Nish S., Shimamura S., Suzuki H.,
 RA Matsui S., Uchiyama I.;
 RT "Thermoadaptation trait revealed by the genome sequence of
 RT thermophilic Geobacillus kaustophilus.";
 RL Nucleic Acids Res. 32:6292-6303(2004).
 DR EMBL; BA000043; BAD75201.1; -; Genomic_DNA.
 DR InterPro; IPR007544; Linocin_M18.
 DR Pfam; PF04454; Linocin_M18; I.
 KW Complete proteome; Hypothetical protein.
 SQ SEQUENCE 284 AA; 31836 MW; EFCD894846619166 CRC64;

Query Match 14.6%; Score 196.5; DB 2; Length 284;
 Best Local Similarity 26.1%; Pred. No. 9.4e-07;
 Matches 68; Conservative 49; Mismatches 117; Indels 27; Gaps 11;

QY 4 LYDLDLPVTEAANAETLEAAKTFKRIAGRVVDVSDPGPVTAAVSTGRIDVKAFTN 63
 DB 6 LYPE-APLTSSQNGEDELVIETARQLVGRFPLDGLPGEGVGVANDIYN--PEQ 61
 QY 64 GVLA----HLRASKPLVRLRVP-----FTLSRNEIDVVERGSKDSDFEVPKEAKKLA 112
 DB 62 GDSFTGCKELSLSEPARRVLTTPLYKDPILTYRDLKQKQKQSPIDFSAANAAQCCA 121
 QY 113 FVEDRTTFEGYSAASIEGIRASASPALTLP--DPRRIPDVISQALSELRLAGVDGPY 169
 DB 122 LLEDLDLIFNGSTFEDVGINNVKGIHRSQWKSNAFTDIV-ELRNKLLDQGHGTPY 180
 QY 170 SVLISADVTKVSTSHGVPPIR-ELHRLVVDGDIYAPLIDG-AFVLTTRGGDPPLQL 226
 DB 181 ALVLSPELVALLIRHVEGTHVLEIHEIRLMTAGIYOTPVYIKGKGVVITGRQNDLAV 240
 QY 227 GTDVAIGYASHDTERLT-YL 246
 DB 241 AVDVQTAFA---LDTEMMNTYL 257

RESULT 13
 QANSX8_9DELTA PRELIMINARY; PRT; 272 AA.
 AC QANSX8;
 DT 13-SEP-2005 (TrEMBLrel. 31, Created)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
 DE Similar to Uncharacterized protein linocin/Cpf29 homolog.
 GN ORFNames=AdendRAFT_1692;
 OS Anaeromyxobacter dehalogenans 2CP-C.
 OC Bacteria; Proteobacteria; Deltaproteobacteria; Myxococcales;
 OC Cyetobacterineae; Myxococcaceae; Anaeromyxobacter.
 NCBI_TaxID=290397;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=2CP-C;
 RG US DOE Joint Genome Institute (JGI-GenF);
 RA Copeland A., Lucas S., Lapidus A., Barry K., Detter C., Glavina T.,
 RA Hammon N., Israni S., Pitluck S., Richardson P.;
 RT "Sequencing of the draft genome assembly of Anaeromyxobacter
 RT dehalogenans 2CP-C.";
 RN Submitted (May-2005) to the EMBL/GenBank/DBJ databases.

RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=2CP-C;
 RG US DOE Joint Genome Institute (JGI-ORNL);
 RA Larimer F., Land M.;
 RT "Annotation of the draft genome assembly of Anaeromyxobacter
 RT dehalogenans 2CP-C.";
 RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
 CC CAUTION: The sequence shown here is derived from an
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 CC preliminary data.
 DR EMBL; AAHD0100020; BAI78653.1; -; Genomic DNA.
 SQ SEQUENCE 272 AA; 28470 MW; 477948F0A22C0A49 CRC64;

Query Match 13.9%; Score 187.5; DB 2; Length 272;
 Best Local Similarity 28.2%; Pred. No. 4.3e-06;
 Matches 78; Conservative 40; Mismatches 128; Indels 31; Gaps 11;

QY 6 RDLAPVTEAANAETLEAAKTFKRIAGRVVDVSDPGPVTAAVSTGRIDVKAFTN 62
 DB 6 RDLAPVTEAANAETLEAAKTFKRIAGRVVDVSDPGPVTAAVSTGRIDVKAFTN 62
 QY 63 NG-----VIAHLRASKPLVRLRVPFTLSRNEIDVVERGSKDSDFEVPKEAKKLA 113
 DB 58 GGEDEPBADETHVVPVRLPVLHRTFRLGAAVBALERGEPVLTTEAANAARIRAR 117
 QY 114 VEDRTTFEGYSAASIEGIRASASPALTLP--DPRRIPDVISQALSELRLAGVDGPY 168
 DB 118 AEDRLTFEGHAGA---GVRLLEHPEGLVEVPAGDMADPGRAGDALLAALTDAGRHGP 174
 QY 169 YSVLS-ADVTKVSTSHGVPPIR-ELHRLVVDGDIYAPLIDG-AFVLTTRGGDPPLQL 226
 DB 175 YAAVSPAPRFYQFRFPAGTALTPYQQLPAREGGVVKAPGLRDGAVVVRSASGQAVV 234
 QY 227 GTDVAIGYASHDTERLT-YL 263
 DB 235 GDELTAAYGREGIFHLVSLASVTLL-PGAPGSVAV 270

RESULT 14
 067639_AQUAE PRELIMINARY; PRT; 281 AA.
 AC 067639;
 DT 01-AUG-1998 (TrEMBLrel. 07, Created)
 DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Hypothetical protein aq_1760.
 GN OrderedLocuNames=AQ_1760;
 OS Aquifex aeolicus.
 OC Bacteria; Aquificae; Aquificales; Aquificaceae; Aquifex.
 NCBI_TaxID=63363;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=VPS;
 RX MEDLINE=98196666; PubMed=9537320; DOI=10.1038/32831.
 RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
 RA Graham D.E., Overbeek R., Sneed M.A., Keller M., Aulay M., Huber R.,
 RA Feldman R.A., Short J.M., Olsen G.J., Swanson R.V.;
 RT "The complete genome of the hyperthermophilic bacterium Aquifex
 RT aeolicus.";
 RL Nature 392:353-358(1998).
 DR EMBL; AB000754; AAC07600.1; -; Genomic_DNA.
 DR PIR; E70451; E70451.
 DR InterPro; IPR007544; Linocin_M18.
 DR Pfam; PF04454; Linocin_M18; I.
 KW Complete proteome.
 SQ SEQUENCE 281 AA; 31847 MW; C0B36F6296801631 CRC64;

Query Match 10.5%; Score 142; DB 2; Length 281;
 Best Local Similarity 24.2%; Pred. No. 0.013;
 Matches 63; Conservative 41; Mismatches 120; Indels 36; Gaps 9;
 QY 1 MNLYDLDLPVTEAANAETLEAAKTFKRIAGRVVDVSDPGPVTAAVSTGRIDVKAFTN 45

```

Db      1 MBFLQDQAPLTAEBWEQIDKTAVEVFKSTVVCRRKFMVGPFGAGHOVSYDVLVYGP 60
Qy      46 VTAAVSTGRLLDVAKPT-NGVIAHLBASKPLVRLRPPTLSRNEIDIDVERGSKDSMEPV 104
        |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db      61 GVCVKKQCGQBYKVCBPARTGRKIV----PVPTLYKQFVISMRLDHRKQNLVPDITGV 116
        |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Qy      105 KEAAKQLAFVEDRTTFEGYSAAISIGIRSSASNPALTLPEDPRE-----IPVISOAL 157
        |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db      117 AAAASSIAVADKLLIFGNQEMGIEGFLLTAG-----TLREELSDMEKVGNAFOVIV-KGI 171
        |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Qy      158 SELRLAGVDGYSVLLSADVYTKVSETSDH-GYPIREHLANLNDGDIIMAPAL--DGAFAV 214
        |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db      172 SRLVEKPEPTNYIYLIVPKRYPFLINRHHDNGLLELQIKKVKV-EVYQTEIIPEDIVLL 230
        |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Qy      215 LITRGDPDLQLTGTDVAIGY 234
        |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db      231 VSASPANFDLAIADVNVAF 250
        |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

```

RESULT 15

```

026970_TRYCR
ID 026970_TRYCR PRELIMINARY; PRT; 453 AA.
AC 026970;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)
DE Trypanothione reductase (Fragment).
OS Trypanosoma cruzi.
OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma;
OC Schizotrypanum.
OX NCBI_TaxID=5693;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CAI;
RX MEDLINE=96320199; PubMed=8702014;
RA Nozaki T., Engel J.C., Dvorak J.A.;
RT "Cellular and molecular biological analyses of nitfurimox resistance
   in Trypanosoma cruzi.";
RL Am. J. Trop. Med. Hyg. 55:111-117(1996).
CC -I- SIMILARITY: Belongs to the class-I pyridine nucleotide-disulfide
   oxidoreductase family.
CC EMBL, M97953; AAB59211.1; -; Genomic_DNA.
DR PDB; 1BZL; X-ray; A/B=1-453.
DR GO; GO:0005737; C:cytoplasm; IEA.
DR GO; GO:0050660; F:disulfide oxidoreductase activity; IEA.
DR GO; GO:0006118; F:electron transport; IEA.
DR InterPro; IPR013271; FAD_pyr_redox.
DR InterPro; IPR004099; Pyr_redox_dlm.
DR InterPro; IPR001864; Trypanth_redctse.
DR Pfam; PF00070; Pyr_redox; 1.
DR Pfam; PF02852; Pyr_redox_dlm; 1.
DR PRINTS; PR00368; FADPNR.
DR PRINTS; PR00411; PNDRDTASRI.
DR PRINTS; PR00470; TRYPARNDTASR.
DR TIGRPFAMs; TIGR01423; trypano_reduc; 1.
DR PROSITE; PS00076; PYRIDINE_REDOX_1; 1.
KW FAD; Flavoprotein; Oxidoreductase; Redox-active center.
FT NON_TER 1
FT NON_TER 453
SQ SEQUENCE 453 AA; 49416 MW; 6BD6079B62AF51B CRC64;

```

Query Match 8.3%; Score 112; DB 2; Length 453;

Best Local Similarity 25.8%; Pred. No. 4.9; Matches 96; Indels 82; Gaps 18;

```

Qy      13 EAANAEIELEAARTFKHIAGRVVDSDPGPYTAAVSTGRLLDV-----KAPTNQV-- 65
        |||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db      11 EAAN-----NDAITYKKKVA---VIDVOMVHGPPFPFSLGTCVAVGCVPKGLMVTGAQY 62
        |||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Qy      66 IAHLRAS-----KPLVLRVVFLLSRNE-IDDYER-----GSK 97
        |||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

```

```

Db      63 MEHLRESAGGWEFDRITTLBAEMKLLIAYDCAVLNINKSYDENFRDTEGLEFLGMSL 122
Qy      98 DS-DMEPVKAARKLAFVEDRTTFEGYSAA-----IEGIRSA-SSNPALTLPEDPRE 147
        |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db      123 ESKVAVVVRSSADPASAVKRLTEHLLASGSWPHPNIPGIEHCISNNAFLLPEPFR 182
        |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Qy      148 EIPVISOALSELRLAGV-----DGPYSVLLSADVYTKVSETSDHGYPIREHLANLND 200
        |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db      183 RVLTVGGGFIS-VEFAGIFNAVYKPKDQVTLCTYRGEH---ILRGFDH--TLREELTKQLT 236
        |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Qy      201 GDITWA-----IDGAFVLTTRG---DEFLQLGTDVAIGVASHDTDE 242
        |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db      237 ANGIQILTKENPAKVBINADGSKSVYESGKQDFLVM--MAGKSPRTKDQ 288
        |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

```

Search completed: April 14, 2006, 17:32:09
Job time : 101.776 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using SW model

Run on: April 14, 2006, 17:32:59 ; Search time 24.5038 Seconds
(without alignments)
894.107 Million cell updates/sec

Title: US-10-620-246-16

Perfect score: 1348
Sequence: 1 MNNLYRDLAPVTEAAMAEIR.....LOSTLPLCYTAASVALSH 265

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA:*
1: /cgn2_6/ptodata/1/1aa/5_COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/6_COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/H_COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/PC/US_COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/RE_COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1348	100.0	265	US-09-050-739-16	Sequence 16, Appl
2	147	10.9	287	US-09-902-540-15099	Sequence 15099, A
3	100	7.4	19	US-09-050-739-23	Sequence 23, Appl
4	94.5	7.0	853	US-09-354-129-10	Sequence 10, Appl
5	94.5	7.0	853	US-09-504-357-10	Sequence 10, Appl
6	93.5	6.9	269	US-09-252-991A-20218	Sequence 20218, A
7	91	6.8	311	US-09-489-039A-8355	Sequence 8355, Ap
8	90	6.7	530	US-09-252-991A-20287	Sequence 20287, A
9	89.5	6.6	268	US-09-252-991A-32791	Sequence 32791, A
10	89.5	6.6	755	US-09-252-991A-18091	Sequence 18091, A
11	88.5	6.6	744	US-09-252-991A-29565	Sequence 29565, A
12	87.5	6.5	538	US-09-461-774-2	Sequence 2, Appl1
13	86.5	6.4	558	US-09-071-035-268	Sequence 268, App
14	86.5	6.4	558	US-10-206-576-268	Sequence 268, App
15	86.5	6.4	1638	US-09-071-035-258	Sequence 258, App
16	86.5	6.4	1638	US-09-071-035-262	Sequence 262, App
17	86.5	6.4	1638	US-09-071-035-266	Sequence 266, App
18	86.5	6.4	1638	US-10-206-576-258	Sequence 258, App
19	86.5	6.4	1638	US-10-206-576-262	Sequence 262, App
20	86.5	6.4	1638	US-10-206-576-266	Sequence 266, App
21	86.5	6.4	1747	US-09-134-000C-55999	Sequence 55999, Ap
22	86	6.4	338	US-08-961-536-2	Sequence 2, Appl1
23	86	6.4	338	US-09-583-110-4275	Sequence 4275, Ap
24	86	6.4	338	US-09-432-682-2	Sequence 2, Appl1
25	86	6.4	346	US-09-107-433-4305	Sequence 4305, Ap
26	86	6.4	382	US-09-252-991A-29696	Sequence 29696, A
27	86	6.4	1163	US-09-902-540-10418	Sequence 10418, A

28	86	6.4	1601	2	US-09-345-473B-40	Sequence 40, Appl
29	86	6.4	1601	2	US-09-862-027-40	Sequence 40, Appl
30	85.5	6.3	1287	2	US-09-252-991A-29606	Sequence 29606, A
31	85	6.3	330	2	US-09-583-110-3623	Sequence 3623, Ap
32	85	6.3	2216	2	US-09-902-540-12221	Sequence 12221, A
33	85	6.3	4572	2	US-10-042-665A-4	Sequence 4, Appl1
34	84.5	6.3	297	2	US-09-605-703B-296	Sequence 296, App
35	84.5	6.3	310	2	US-09-252-991A-31483	Sequence 31483, A
36	84.5	6.3	828	2	US-09-605-703B-294	Sequence 294, App
37	84.5	6.3	313	2	US-09-489-039A-12219	Sequence 12219, A
38	84	6.2	222	2	US-09-107-532A-4165	Sequence 4165, Ap
39	84	6.2	280	2	US-09-902-540-12020	Sequence 12020, A
40	84	6.2	377	2	US-09-902-540-16656	Sequence 16656, A
41	84	6.2	1820	2	US-09-902-540-16307	Sequence 16307, A
42	83.5	6.2	225	2	US-09-248-335-30	Sequence 30, Appl
43	83.5	6.2	1525	2	US-09-396-651B-1	Sequence 1, Appl1
44	83.5	6.2	1721	2	US-10-042-665A-6	Sequence 6, Appl1
45	83	6.2	444	2	US-09-328-352-7043	Sequence 7043, Ap

ALIGNMENTS

RESULT 1						
US-09-050-739-16						
Sequence 16, Application US/09050739						
Patent No. 6641814						
GENERAL INFORMATION:						
APPLICANT: ANDERSEN, Peter						
APPLICANT: NIELSEN, Rikke						
APPLICANT: OETTINGER, Thomas						
APPLICANT: RASMUSSEN, Peter Birk						
APPLICANT: ROSENKRANDS, Ida						
APPLICANT: WELDMING, Karin						
APPLICANT: FLORIO, Walter						
TITLE OF INVENTION: NUCLEIC ACIDS FRAGMENTS AND POLYPEPTIDE FRAGMENTS						
FILE REFERENCE: 670001-2002.1						
CURRENT APPLICATION NUMBER: US/09/050,739						
CURRENT FILING DATE: 1998-03-30						
EARLIER APPLICATION NUMBER: 0376/97						
EARLIER FILING DATE: 1997-04-02						
EARLIER APPLICATION NUMBER: 1277/97						
EARLIER FILING DATE: 1997-11-10						
EARLIER APPLICATION NUMBER: 60/044,624						
EARLIER FILING DATE: 1997-04-18						
EARLIER APPLICATION NUMBER: 60/070,488						
EARLIER FILING DATE: 1998-01-05						
NUMBER OF SEQ ID NOS: 173						
SOFTWARE: Patentin Ver. 2.0						
SEQ ID NO 16						
LENGTH: 265						
TYPE: PRT						
ORGANISM: Mycobacterium tuberculosis						
US-09-050-739-16						
Query Match						
Best Local Similarity 100.0%; Score 1348; DB 2; Length 265;						
Matches 265; Conservative 0; Mismatches 0; Indels 0; Gaps 0;						
QY	1	MNNLYRDLAPVTEAAMAEIR	ELBAARTEFKRIAGRRVVDSPGPGVTAAGVSTGLIDVKA	60		
DB	1	MNNLYRDLAPVTEAAMAEIR	ELBAARTEFKRIAGRRVVDSPGPGVTAAGVSTGLIDVKA	60		
QY	61	PTNGVIAHLASRPLVRLRVPFTLSRNEIDVBRGSDSWEPYKAAKLAFTYEDRTIF	120			
DB	61	PTNGVIAHLASRPLVRLRVPFTLSRNEIDVBRGSDSWEPYKAAKLAFTYEDRTIF	120			
QY	121	EGSASASIRGIRASASNPALTLPEDPREIPDIVISQALSERLAGVGPYSVLISADVYTK	180			
DB	121	EGSASASIRGIRASASNPALTLPEDPREIPDIVISQALSERLAGVGPYSVLISADVYTK	180			
QY	181	VSETSDHGYPRIREHLNRLVVDGIYMAPALDGAFLVLTTRGGDFDLQCTDVAIGYASHDTD	240			

1 CURRENT APPLICATION NUMBER: US/09/504,357
2 ;
3 ; CURRENT FILING DATE: 2000-02-15
4 ; EARLIER APPLICATION NUMBER: 60/093,181
5 ; EARLIER FILING DATE: July 17, 1998
6 ; NUMBER OF SEQ ID NOS: 14
7 ; SOFTWARE: Microsoft Office 97
8 ; SEQ ID NO 10
9 ; LENGTH: 853
10 ; TYPE: PRT
11 ; ORGANISM: Pseudomonas sp. strain G-179
12 US-09-504-357-10

Query Match 7.0%; Score 94.5; DB 2; Length 853;
Best Local Similarity 25.1%; Pred. No. 0.5;
Matches 59; Conservative 27; Mismatches 74; Indels 75; Gaps 15;

77 RLAVPPLSLRNEIDVGRSGKSDMEPV-----KEAKKLAFFEDR----- 117
119 RLTPPLLRKNGVFD-----KEGFEFVWEBAFDIAEAKAKT--LKEKGPTALMGFGS 171
118 ---TIFEGYASASI--EGIRASASNPALTLPEDPREIPDVISQALSELRLAGVDP---Y 169
172 GQMTIFEGYATATLMRAGFRSNNL-----DPNAHCHASAAAGMTIFGMDPEMGCT 223
170 SVLLSADVTYKVBSTSDHGYPIREHLNRLVVDGIWAPADGAF-----VLT---R 218
224 DDEHADAFVLMGSMNAEMKRP-----ILMTRLABDRLGHEHVKVSVLSTPTHR 271
219 GGP-GLQ-----GTDAI-GY-ASHPTDTERL---YLQETLPLCTAASVAL 263
272 SMDLADIPLVFKDGTDLALNYLANHIIQGRVNOQEPIDKNTKMOATTDIGYGL 326

RESULT 6

US-09-252-991A-20218
1 ; Sequence 20218; Application US/09252991A
2 ; Patent No. 6551795
3 ; GENERAL INFORMATION:
4 ; APPLICANT: Marc J. Rubenfield et al.
5 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
6 ; FILE REFERENCE: 107196.136
7 ; CURRENT APPLICATION NUMBER: US/09/252,991A
8 ; CURRENT FILING DATE: 1999-02-18
9 ; PRIOR APPLICATION NUMBER: US 60/074,788
10 ; PRIOR FILING DATE: 1998-02-18
11 ; PRIOR APPLICATION NUMBER: US 60/094,190
12 ; NUMBER OF SEQ ID NOS: 33142
13 ; SEQ ID NO 20218
14 ; LENGTH: 269
15 ; TYPE: PRT
16 ; ORGANISM: Pseudomonas aeruginosa
17 US-09-252-991A-20218

Query Match 6.9%; Score 93.5; DB 2; Length 269;
Best Local Similarity 22.7%; Pred. No. 0.096;
Matches 57; Conservative 28; Mismatches 91; Indels 75; Gaps 11;

30 HIAGRVVDVSDGPGVTAAVSTGRLLDVKAPTNNGVIAHLRASKPLVRLVPTLSRNEI 89
7 HHAGAR-----PDRAPVA-----LRSARAPRHASALQRADRA-----TLRQGA 48
90 DDERGSKSDMWPV-----KEAKKLAFFEDRITIFEGYASASIEGIRASASNP 138
49 DPTMDPATAD-BPVRLLVVDHPMMKGVACLELEDDLSVVGESGSEBALRLA--- 104
139 ALTLPRDPREI-PPVI-----SQALSELRLAGVDPYVLLSAD---VYTVKSE 183
105 -----ELDDPMILIDLMKMGANGLDTLRALREGVAGARLVFVVSDDKGDVAVVLR 155
184 TSDHGYPIR-----EHLNRLVVDGIWAPADGAFVLTTRGSD-----FDLQ 225

Db 156 AGADGILLKMEPERLLEHITROAATGOMTSLPOLTOILAQALRGDRSKSLDELTERBQ 215
Cy 226 LGTDVAIGYAS 236
Db 216 ILRQIAHGYSN 226

RESULT 7

US-09-489-039A-8355
1 ; Sequence 8355; Application US/09489039A
2 ; Patent No. 6610835
3 ; GENERAL INFORMATION:
4 ; APPLICANT: Gary Breton et. al
5 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
6 ; FILE REFERENCE: 2709.2004001
7 ; CURRENT APPLICATION NUMBER: US/09/489,039A
8 ; CURRENT FILING DATE: 2000-01-27
9 ; PRIOR APPLICATION NUMBER: US 60/117,747
10 ; PRIOR FILING DATE: 1999-01-29
11 ; NUMBER OF SEQ ID NOS: 14342
12 ; SEQ ID NO 8355
13 ; LENGTH: 381
14 ; TYPE: PRT
15 ; ORGANISM: Klebsiella pneumoniae
16 US-09-489-039A-8355

Query Match 6.8%; Score 91; DB 2; Length 381;
Best Local Similarity 23.1%; Pred. No. 0.32;
Matches 28; Conservative 32; Mismatches 51; Indels 10; Gaps 4;

84 LSRNEID-----VERGSKSDMEPVKEAKKLAFFEDRTIF--EGYASASIEGIRASAS 135
96 IARNKINNNSYLLFVDDGSRDNTMAQIKOASTAYHYRGIKLSRNKGHOIALMAGLRSD 155
Cy 136 SNPALTLPEDPREIPDVISQALSELRLAGVDPYVLLSADVTYKVBSTSDHGYPIREH 194
Db 156 TDVTISIDADLQDVCIEKMI-DAYSQGYDIYGVRCNRDSDTPFRRTTANAFYALMSH 214
Cy 195 L 195
Db 215 W 215

RESULT 8

US-09-252-991A-20287
1 ; Sequence 20287; Application US/09252991A
2 ; Patent No. 6551795
3 ; GENERAL INFORMATION:
4 ; APPLICANT: Marc J. Rubenfield et al.
5 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
6 ; FILE REFERENCE: 107196.136
7 ; CURRENT APPLICATION NUMBER: US/09/252,991A
8 ; CURRENT FILING DATE: 1999-02-18
9 ; PRIOR APPLICATION NUMBER: US 60/074,788
10 ; PRIOR FILING DATE: 1998-02-18
11 ; PRIOR APPLICATION NUMBER: US 60/094,190
12 ; NUMBER OF SEQ ID NOS: 33142
13 ; SEQ ID NO 20287
14 ; LENGTH: 530
15 ; TYPE: PRT
16 ; ORGANISM: Pseudomonas aeruginosa
17 US-09-252-991A-20287

Query Match 6.7%; Score 90; DB 2; Length 530;
Best Local Similarity 23.5%; Pred. No. 0.72;
Matches 54; Conservative 32; Mismatches 78; Indels 66; Gaps 10;

Cy 17 AETLE-----AATFKHIAGRVVDVSDGPGVTAAVSTGRLLDVKAPTNNGVI 66
Db 273 ABRDLVPPVNASQALSGRIGQGEVGRRL-----ALGNRRLLDBQLKPGAL 320

QY 228 TDVAIGYASHDTERLY 245
 DB 304 TANTOGIADGAGTNNRY 321

RESULT 12

US-09-461-774-2
 ; Sequence 2, Application US/09461774
 ; Patent No. 6887481
 ; GENERAL INFORMATION:
 ; APPLICANT: CHAN, Lily
 ; APPLICANT: CHONG, Maxey Ching Ming
 ; APPLICANT: LIM, Renee Lay Hong
 ; TITLE OF INVENTION: Bacterial-derived molecules and therapeutic and
 ; TITLE OF INVENTION: diagnostic uses thereof
 ; FILE REFERENCE: 1781-0180P
 ; CURRENT APPLICATION NUMBER: US/09/461,774
 ; CURRENT FILING DATE: 1999-12-15
 ; NUMBER OF SEQ ID NOS: 22
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 2
 ; LENGTH: 538
 ; TYPE: PRT
 ; ORGANISM: Mycobacterium tuberculosis
 ; US-09-461-774-2

Query Match 6.5%; Score 87.5; DB 2; Length 538;
 Best Local Similarity 25.2%; Pred. No. 1.4; Indels 99; Gaps 13;
 Matches 66; Conservative 19; Mismatches 78; Indels 99; Gaps 13;

QY 7 DLAPVTEA-----AMAEIELEAPRTFKRIAGRRVVDSDPGPVTAAVSTGRLLID 57
 DB 288 DLAVTVGVHVNPDAGIVLRGVLGSGARRVVSDDTVIYDGGTAFAVANRAN--- 344
 QY 58 VKAPVTGVAHMAASRPLVRLVPTLSRNEIDVERGSKDSM--EPVKAACKLA--- 112
 DB 345 -----HLRA-----EID-----KSDSDMDREKLGRLAKLAGV 373
 QY 113 -----FYEDRT-----IFEGYSAASIEGIRASNSPALTLPE 144
 DB 374 AVIKGAADDTALKERKESVEDAVAAKAAVEGIVPGGASLIHQARKALLTLRSLTG 433
 QY 145 DPREIDVISOALSE-----LRLAGVDGPPYVLLSADVTVKVS--TSDGYPYRHLNRL 198
 DB 434 DRYLGVDVSEALAAPLFWIAANAGLDG-----SVVNAKVSLELPRGHL-----INVTL 482
 QY 199 VDGDIIMAPADGAF--VLTR 218
 DB 483 SYGDL-----AADGVIDPVKVT 500

RESULT 13

US-09-071-035-268
 ; Sequence 268, Application US/09071035
 ; Patent No. 6448043
 ; GENERAL INFORMATION:
 ; APPLICANT: Gil H. Choi
 ; TITLE OF INVENTION: Enterococcus faecalis polynucleotides and polypeptides
 ; NUMBER OF SEQUENCES: 496
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSER: Human Genome Sciences, Inc.
 ; STREET: 9410 Key West Avenue
 ; CITY: Rockville
 ; STATE: Maryland
 ; COUNTRY: USA
 ; ZIP: 20850
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
 ; OPERATING SYSTEM: MSDOS version 6.2
 ; SOFTWARE: ASCII Text
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/071,035

; FILING DATE:
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER:
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: A. Anders Brookes
 ; REGISTRATION NUMBER: 36,373
 ; REFERENCE/DOCKET NUMBER: PB369P2
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (301) 309-8504
 ; TELEFAX: (301) 309-8512
 ; INFORMATION FOR SEQ ID NO: 268:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 558 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-09-071-035-268

Query Match 6.4%; Score 86.5; DB 2; Length 558;
 Best Local Similarity 24.4%; Pred. No. 1.9; Indels 73; Gaps 13;
 Matches 64; Conservative 28; Mismatches 97; Indels 73; Gaps 13;

QY 37 VDVSDFGPVTAIVSTG-----RLIDVAPFNGVY-----AHLASRPLVRLVPT 83
 DB 139 VEVRYDADGVLMGATGTGDKTVTLERPGASANEITVAKNATGESOP-ATATTPVD 197
 QY 84 LSRNEIDVERGSKDSDEPVKAACKLAF--VEDRTIEGYSAASIEGIRASNSPAL 140
 DB 198 LATPTIDSI-TGNSKGEITGYABPKTTIDVADADGTIIAATTA-----NETQYTV 249
 QY 141 TLPE---DPREIDVSO--ALSELRLAGVDGPPYVLLSADVTVKVSSETHGPI--- 191
 DB 250 TLPAVVTVGETTITLSKAGNESQPAIVADAVLAAPTITKVEGNANGTVTGTA 309
 QY 192 -----REHLRLVVDGDIIMAPADGAFVLTTRGSDFDLOLGTDVAVIGASHDT 240
 DB 310 DRYVTVQFYNSSQL--LASGN-----TTGGTFVHIAAGLA----- 345
 QY 241 TERLYLQETITFLCYTAASVA 262
 DB 346 TER-----ETTLALTDTQGNVS 363

RESULT 14

US-10-206-576-268
 ; Sequence 268, Application US/10206576
 ; Patent No. 6913907
 ; GENERAL INFORMATION:
 ; APPLICANT: Choi et al.
 ; TITLE OF INVENTION: Enterococcus faecalis polynucleotides and polypeptides
 ; NUMBER OF SEQUENCES: 497
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSER: Human Genome Sciences, Inc.
 ; STREET: 9410 Key West Avenue
 ; CITY: Rockville
 ; STATE: Maryland
 ; COUNTRY: USA
 ; ZIP: 20850
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: CD-R
 ; OPERATING SYSTEM: Windows 98
 ; SOFTWARE: ASCII Text
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/10/206,576
 ; FILING DATE: 29-Jul-2002
 ; CLASSIFICATION: <Unknown>
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 09/071,035
 ; FILING DATE: 1998-05-04

APPLICATION NUMBER: US 60/046,655
FILING DATE: 1997-05-16
APPLICATION NUMBER: US 60/044,031
FILING DATE: 1997-05-06
APPLICATION NUMBER: US 60/066,009
FILING DATE: 1997-11-14
ATTORNEY/AGENT INFORMATION:
NAME: Hyman, Mark J.
REGISTRATION NUMBER: 46,789
REFERENCE/DOCKET NUMBER: PB369P1D1
INFORMATION FOR SEQ ID NO: 268:
SEQUENCE CHARACTERISTICS:
LENGTH: 558 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 268:
US-10-206-576-268

Query Match 6.4%; Score 86.5; DB 2; Length 558;
Best Local Similarity 24.4%; Pred. No. 1.9; Mismatches 97; Indels 73; Gaps 13;
Matches 64; Conservative 28; Mismatches 97; Indels 73; Gaps 13;

QY 37 VDVSDEGPGVTAAVSTG-----RLIDVKAPTNGVI-----AHLRASKPLVRLRVPT 83
DB 139 VEVRDADGTVLGWATTGTDGKTVTLPEPKASANETITVAKNATGESQP-ATATTPVD 197
QY 84 LSRNEIDYVERGSKDSMDPEVKEAKKLAP---VEDRTIFEGYSASISIRASSNPAL 140
DB 198 LAPPTIDSI-TGSSSKGYEITGTAEPKTTIDVRDADGTTIAATTA-----NETQYTV 249
QY 141 TLPE---DPRRIDVISO--ALSELRLAGVDGPFYSVLSADVYTKVSESDHGYP----- 191
DB 250 TLDPAGVVTGEGTTITISKDAGNESQPAITAVIPADVLAAPITTKVKGKANGYTTGTGA 309
QY 192 -----REHLNRLVGDIIWAPADGAFVLTTRGSDPDLQGTGTVAGYASHDT 240
DB 310 DPNVTQGFYNSSQL--LASGN-----TTTGGTFSVHIAAGLA----- 345
QY 241 TERLYAOETITFLCTAASVA 262
DB 346 TERK---ETLTALTDTDQGNVS 363

RESULT 15

US-09-071-035-258
Sequence 258, Application US/09071035
Patent No. 6448043
GENERAL INFORMATION:
APPLICANT: Gil H. Choi
TITLE OR INVENTION: Enterococcus faecalis polynucleotides and polypeptides
NUMBER OF SEQUENCES: 496
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/071,035
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:

NAME: A. Anders Brookes
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB369P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 258:
SEQUENCE CHARACTERISTICS:
LENGTH: 1638 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-071-035-258

Query Match 6.4%; Score 86.5; DB 2; Length 1638;
Best Local Similarity 24.4%; Pred. No. 11; Mismatches 97; Indels 73; Gaps 13;
Matches 64; Conservative 28; Mismatches 97; Indels 73; Gaps 13;

QY 37 VDVSDEGPGVTAAVSTG-----RLIDVKAPTNGVI-----AHLRASKPLVRLRVPT 83
DB 1184 VEVRDADGTVLGWATTGTDGKTVTLPEPKASANETITVAKNATGESQP-ATATTPVD 1242
QY 84 LSRNEIDYVERGSKDSMDPEVKEAKKLAP---VEDRTIFEGYSASISIRASSNPAL 140
DB 1243 LAPPTIDSI-TGSSSKGYEITGTAEPKTTIDVRDADGTTIAATTA-----NETQYTV 1294
QY 141 TLPE---DPRRIDVISO--ALSELRLAGVDGPFYSVLSADVYTKVSESDHGYP----- 191
DB 1295 TLDPAGVVTGEGTTITISKDAGNESQPAITAVIPADVLAAPITTKVKGKANGYTTGTGA 1354
QY 192 -----REHLNRLVGDIIWAPADGAFVLTTRGSDPDLQGTGTVAGYASHDT 240
DB 1355 DPNVTQGFYNSSQL--LASGN-----TTTGGTFSVHIAAGLA----- 1390
QY 241 TERLYAOETITFLCTAASVA 262
DB 1391 TERK---ETLTALTDTDQGNVS 1408

Search completed: April 14, 2006, 17:37:52
Job time : 26.5038 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: April 14, 2006, 18:27:13 ; Search time 92.995 Seconds
(without alignments)
1190.652 Million cell updates/sec

Title: US-10-620-246-16

Perfect score: 1348

Sequence: 1 MNNTLYRDLAPVTAAABAEI.....LQETTLFLCYTAASVALSH 265

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA Main:
1: /cgn2_6/ptodata/1/pubppaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubppaa/US08_PUBCOMB.pep:*
3: /cgn2_6/ptodata/1/pubppaa/US09_PUBCOMB.pep:*
4: /cgn2_6/ptodata/1/pubppaa/US10_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubppaa/US10_PUBCOMB.pep:*
6: /cgn2_6/ptodata/1/pubppaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1348	100.0	265	3	US-09-791-171-16
2	1348	100.0	265	3	US-09-804-980-16
3	1348	100.0	265	3	US-10-620-246-16
4	108	8.0	1574	4	US-10-205-032-16
5	103	7.6	8805	6	US-11-097-143-28128
6	101	7.5	554	4	US-10-125-692-21
7	101	7.5	554	5	US-10-991-347-21
8	100	7.4	19	3	US-09-791-171-23
9	100	7.4	19	3	US-09-804-980-23
10	100	7.4	19	4	US-10-620-246-23
11	98	7.3	410	4	US-10-156-761-14324
12	98	7.3	410	5	US-10-732-923-1986
13	96	7.1	555	4	US-10-369-493-9172
14	96	7.1	579	4	US-10-369-493-9415
15	96	7.1	614	4	US-10-369-493-17504
16	96	7.0	621	6	US-11-097-143-25428
17	95	7.0	4038	3	US-09-980-217-22
18	95	7.0	4038	5	US-10-732-923-20546
19	92.5	6.9	583	4	US-10-369-566-1542
20	92	6.8	728	4	US-10-425-114-54770
21	91	6.8	472	5	US-10-739-930-9412
22	91	6.8	702	4	US-10-282-122A-52256
23	90.5	6.7	526	4	US-10-369-493-18051
24	90.5	6.7	1220	5	US-10-450-763-54677
25	90	6.7	328	4	US-10-282-122A-71439
26	90	6.7	464	4	US-10-369-493-15296
27	90	6.7	505	4	US-10-369-493-15662

28	90	6.7	505	4	US-10-369-493-15053	Sequence 16053, A
29	90	6.7	732	3	US-09-815-242-11966	Sequence 11966, A
30	90	6.7	792	4	US-10-127-032-165	Sequence 165, App
31	90	6.7	792	4	US-10-282-122A-66574	Sequence 66574, A
32	89.5	6.6	415	4	US-10-437-963-174225	Sequence 174225, A
33	89.5	6.6	416	4	US-10-369-493-4737	Sequence 4737, Ap
34	89.5	6.6	416	4	US-10-369-493-7496	Sequence 7496, Ap
35	89	6.6	309	4	US-10-369-493-660	Sequence 660, App
36	89	6.6	684	4	US-10-369-493-12042	Sequence 12042, A
37	88.5	6.6	299	3	US-09-738-626-4484	Sequence 4484, Ap
38	88.5	6.6	317	4	US-10-156-761-10897	Sequence 10897, A
39	88.5	6.6	342	5	US-10-494-836-112	Sequence 112, App
40	88.5	6.6	488	3	US-09-893-519A-33	Sequence 33, Appl
41	88.5	6.6	751	4	US-10-156-761-8159	Sequence 8159, Ap
42	88.5	6.6	765	5	US-10-732-923-7244	Sequence 7244, Ap
43	88.5	6.6	767	4	US-10-425-115-322796	Sequence 322796, A
44	88.5	6.6	810	4	US-10-425-114-54293	Sequence 54293, A
45	88.5	6.6	1185	4	US-10-437-963-173561	Sequence 173561, A

ALIGNMENTS

RESULT 1
US-09-791-171-16
; Sequence 16, Application US/09791171
; Patent No. US20020094336A1
; GENERAL INFORMATION:
; APPLICANT: ANDERSEN, Peter
; APPLICANT: NIELSEN, Rikke
; APPLICANT: OSTTINGER, Thomas
; APPLICANT: RASMUSSEN, Peter Birk
; APPLICANT: ROSENKRANDS, Ida
; APPLICANT: WEIDING, Karin
; APPLICANT: FLORIO, Walter
; TITLE OF INVENTION: NUCLEIC ACIDS FRAGMENTS AND POLYPEPTIDE FRAGMENTS
; FILE REFERENCE: 670001-2002.1
; CURRENT APPLICATION NUMBER: US/09/791,171
; CURRENT FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 09/050,739
; PRIOR FILING DATE: 1998-03-30
; PRIOR APPLICATION NUMBER: 0376/97
; PRIOR FILING DATE: 1997-04-02
; PRIOR APPLICATION NUMBER: 1277/97
; PRIOR FILING DATE: 1997-11-10
; PRIOR APPLICATION NUMBER: 60/044,624
; PRIOR FILING DATE: 1997-04-18
; PRIOR APPLICATION NUMBER: 60/070,488
; PRIOR FILING DATE: 1998-01-05
; NUMBER OF SEQ ID NOS: 173
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 16
; LENGTH: 265
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-09-791-171-16
Query Match 100.0%; Score 1348; DB 3; Length 265;
Best Local Similarity 100.0%; Pred. No. 1.7e-124;
Matches 265; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MNNTLYRDLAPVTAAABAEI...LQETTLFLCYTAASVALSH 60
DB 1 MNNTLYRDLAPVTAAABAEI...LQETTLFLCYTAASVALSH 60
QY 61 PTNGVIAHLRASXPVLRVFTLSNKEIDVRSKSDSWEPKKAKKLAFVEDRTIF 120
DB 61 PTNGVIAHLRASXPVLRVFTLSNKEIDVRSKSDSWEPKKAKKLAFVEDRTIF 120
QY 121 EGYSAASISIRGIRASNSPALTLPEDPREIPDVISQALSEIRLAGVDCPYVLLSADVYTK 180
DB 121 EGYSAASISIRGIRASNSPALTLPEDPREIPDVISQALSEIRLAGVDCPYVLLSADVYTK 180

[illegible]


```

; PRIOR APPLICATION NUMBER: 09/791,171
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 09/415,884
; PRIOR FILING DATE: 1999-10-08
; PRIOR APPLICATION NUMBER: 60/116,673
; PRIOR FILING DATE: 1999-01-21
; PRIOR APPLICATION NUMBER: 1281/98
; PRIOR FILING DATE: 1998-10-08
; NUMBER OF SEQ ID NOS: 173
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 23
; LENGTH: 19
; TYPE: PR1
; ORGANISM: Mycobacterium tuberculosis
US-10-620-246-23

Query Match          7.4%; Score 100; DB 4; Length 19;
Best Local Similarity 100.0%; Pred. No. 0.0075;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MNNLYRDLAPVTEAMAEI 19
        |||||
Db      1 MNNLYRDLAPVTEAMAEI 19

RESULT 11
US-10-156-761-14324
; Sequence 14324, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMODA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 14324
; LENGTH: 410
; TYPE: PR1
; ORGANISM: Streptomyces avermitilis
US-10-156-761-14324

Query Match          7.3%; Score 98; DB 4; Length 410;
Best Local Similarity 22.6%; Pred. No. 1.3;
Matches 57; Conservative 29; Mismatches 86; Indels 80; Gaps 11;

QY      31 IAGRRV-----VYSDGGR-----VTAAVSTGRIDVKAPNGVIATLRA 71
        :|||:
Db      185 VRGRRLLAADVYADRLGPRDLALPPHVEYIDAKIYGRMAQEAINNALIEHARQ 244
        :|||:
QY      72 SKPLVRLR--VPFTLSHNEIDVVERGSKDSMBPVKEAAKGLAFVEDRTTFEGYSAASIE 129
        |:|||:
Db      245 GKSVRLKGGDPFVFGG-----MEAGALAEV-----GIPCTVVP 280
        :|||:
QY      130 GIRASASSNPAL-----TLPEDPREIIPDVISQALSEIRLAGVDGPY 169
        :|||:
Db      281 GISSISIVPAAGIPVTHRGVAHEFTVSGHVAPEDDERSLVD--WPAIAKLR-----GTL 333
        :|||:
QY      170 SVLLSADVTKVSET--SDHGVPRIREHLNRLVDGDIYAPALIDGAFVL---TTRGDDP-- 223
        :|||:
Db      334 VILMGVVKIGIKIDTLVAHGKPACTPVALVOEGTTAAQRVDATLATVAETVRAQDVKPP 393
        :|||:
QY      224 --LQGTGTVVAIG 233
```

```

Db      394 AVIYGAVDVG 405

RESULT 12
US-10-732-923-1986
; Sequence 1986, Application US/10732923
; Publication No. US20050108791A1
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C
; CURRENT APPLICATION NUMBER: US/10/732,923
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/310,154
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 1986
; LENGTH: 410
; TYPE: PR1
; ORGANISM: Streptomyces avermitilis MA-4680
US-10-732-923-1986

Query Match          7.3%; Score 98; DB 5; Length 410;
Best Local Similarity 22.6%; Pred. No. 1.3;
Matches 57; Conservative 29; Mismatches 86; Indels 80; Gaps 11;

QY      31 IAGRRV-----VYSDGGR-----VTAAVSTGRIDVKAPNGVIATLRA 71
        :|||:
Db      185 VRGRRLLAADVYADRLGPRDLALPPHVEYIDAKIYGRMAQEAINNALIEHARQ 244
        :|||:
QY      72 SKPLVRLR--VPFTLSHNEIDVVERGSKDSMBPVKEAAKGLAFVEDRTTFEGYSAASIE 129
        |:|||:
Db      245 GKSVRLKGGDPFVFGG-----MEAGALAEV-----GIPCTVVP 280
        :|||:
QY      130 GIRASASSNPAL-----TLPEDPREIIPDVISQALSEIRLAGVDGPY 169
        :|||:
Db      281 GISSISIVPAAGIPVTHRGVAHEFTVSGHVAPEDDERSLVD--WPAIAKLR-----GTL 333
        :|||:
QY      170 SVLLSADVTKVSET--SDHGVPRIREHLNRLVDGDIYAPALIDGAFVL---TTRGDDP-- 223
        :|||:
Db      334 VILMGVVKIGIKIDTLVAHGKPACTPVALVOEGTTAAQRVDATLATVAETVRAQDVKPP 393
        :|||:
QY      224 --LQGTGTVVAIG 233
Db      394 AVIYGAVDVG 405

RESULT 13
US-10-369-493-9172
; Sequence 9172, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 9172
; LENGTH: 565
; TYPE: PR1
; ORGANISM: Xylella fastidiosa
US-10-369-493-9172

Query Match          7.1%; Score 96; DB 4; Length 565;
```

Best Local Similarity 26.0%; Pred. No. 3.3;
Matches 68; Conservative 31; Mismatches 111; Indels 52; Gaps 13;

```
QY 8 LAPVTEAAMAEIEBAARTFKRH---IAGRRVVDV--SDPGPVTAAVSTGRLLDVAKPT 62
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 81 LAPTRELA-----IQVAEAFORVAASISGRVLPVYGGGSGGQALAKRGVHVIYVTF- 134
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
QY 63 NGVIAHL-RASKEPLVLRVPFTLSRNEIDVERGSKDSMEPYEAKKLAFYEDRTIFE 121
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 135 GRVIDHLERGTLDLSEIK---TLVLDADDEMLRMGFIED---VEEVLRLKLPASRQVALFS 188
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
QY 122 GYSAASIEGIRASSNPALTLPEDPREIPDVISOALSELRLAGVDGPY---SVLLSADV 178
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 189 ATMPQPIRRI-----AQTYLQDPLE---VTIATKTTTAAINIRORYMWSGLHKLDAL 237
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
QY 179 TKVSETSDHGYPY-----REHLNRLVNDGDIIMAPADGAF-----VLTTRGDDF 222
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 238 TRILEVETPDAMIIFVTKKATETELAEKIQARGITLTAALINGDMQOQRETTIHQKDGKL 237
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
QY 223 DLQGTDVVAIGYASHDTERL 244
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 298 DILVATDV-----AARGLDVERI 315
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
```

RESULT 14
US-10-369-493-9415
Sequence 9415, Application US/10369493

Publication No. US20030233675A1

GENERAL INFORMATION:

APPLICANT: Cao, Yongwei

APPLICANT: Hinkle, Gregory J.

APPLICANT: Slater, Steven C.

APPLICANT: Goldman, Barry S.

APPLICANT: Chen, Xianfeng

TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF

TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES

FILE REFERENCE: 38-10(52052)B

CURRENT APPLICATION NUMBER: US/10/369,493

PRIOR FILING DATE: 2003-02-28

PRIOR APPLICATION NUMBER: US 60/360,039

PRIOR FILING DATE: 2002-02-21

NUMBER OF SEQ ID NOS: 47374

SEQ ID NO 9415

LENGTH: 579

TYPE: PRT

ORGANISM: Xylella fastidiosa

US-10-369-493-9415

Query Match 7.1%; Score 96; DB 4; Length 579;

Best Local Similarity 26.0%; Pred. No. 3.4;

Matches 68; Conservative 31; Mismatches 111; Indels 52; Gaps 13;

```
QY 8 LAPVTEAAMAEIEBAARTFKRH---IAGRRVVDV--SDPGPVTAAVSTGRLLDVAKPT 62
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 73 LAPTRELA-----IQVAEAFORVAASISGRVLPVYGGGSGGQALAKRGVHVIYVTF- 126
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
QY 63 NGVIAHL-RASKEPLVLRVPFTLSRNEIDVERGSKDSMEPYEAKKLAFYEDRTIFE 121
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 127 GRVIDHLERGTLDLSEIK---TLVLDADDEMLRMGFIED---VEEVLRLKLPASRQVALFS 180
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
QY 122 GYSAASIEGIRASSNPALTLPEDPREIPDVISOALSELRLAGVDGPY---SVLLSADV 178
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 181 ATMPQPIRRI-----AQTYLQDPLE---VTIATKTTTAAINIRORYMWSGLHKLDAL 229
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
QY 179 TKVSETSDHGYPY-----REHLNRLVNDGDIIMAPADGAF-----VLTTRGDDF 222
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 230 TRILEVETPDAMIIFVTKKATETELAEKIQARGITLTAALINGDMQOQRETTIHQKDGKL 289
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
QY 223 DLQGTDVVAIGYASHDTERL 244
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 290 DILVATDV-----AARGLDVERI 307
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
```

RESULT 15
US-10-369-493-17504
Sequence 17504, Application US/10369493

Publication No. US20030233675A1

GENERAL INFORMATION:

APPLICANT: Cao, Yongwei

APPLICANT: Hinkle, Gregory J.

APPLICANT: Slater, Steven C.

APPLICANT: Goldman, Barry S.

APPLICANT: Chen, Xianfeng

TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF

TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES

FILE REFERENCE: 38-10(52052)B

CURRENT APPLICATION NUMBER: US/10/369,493

PRIOR FILING DATE: 2003-02-28

PRIOR APPLICATION NUMBER: US 60/360,039

PRIOR FILING DATE: 2002-02-21

NUMBER OF SEQ ID NOS: 47374

SEQ ID NO 17504

LENGTH: 614

TYPE: PRT

ORGANISM: Xylella fastidiosa

US-10-369-493-17504

Query Match 7.1%; Score 96; DB 4; Length 614;

Best Local Similarity 26.0%; Pred. No. 3.8;

Matches 68; Conservative 31; Mismatches 111; Indels 52; Gaps 13;

```
QY 8 LAPVTEAAMAEIEBAARTFKRH---IAGRRVVDV--SDPGPVTAAVSTGRLLDVAKPT 62
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 89 LAPTRELA-----IQVAEAFORVAASISGRVLPVYGGGSGGQALAKRGVHVIYVTF- 142
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
QY 63 NGVIAHL-RASKEPLVLRVPFTLSRNEIDVERGSKDSMEPYEAKKLAFYEDRTIFE 121
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 143 GRVIDHLERGTLDLSEIK---TLVLDADDEMLRMGFIED---VEEVLRLKLPASRQVALFS 196
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
QY 122 GYSAASIEGIRASSNPALTLPEDPREIPDVISOALSELRLAGVDGPY---SVLLSADV 178
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 197 ATMPQPIRRI-----AQTYLQDPLE---VTIATKTTTAAINIRORYMWSGLHKLDAL 245
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
QY 179 TKVSETSDHGYPY-----REHLNRLVNDGDIIMAPADGAF-----VLTTRGDDF 222
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 246 TRILEVETPDAMIIFVTKKATETELAEKIQARGITLTAALINGDMQOQRETTIHQKDGKL 305
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
QY 223 DLQGTDVVAIGYASHDTERL 244
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 306 DILVATDV-----AARGLDVERI 323
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
```

Search completed: April 14, 2006, 18:40:29
Job time: 95.095 secs

Db 224 AKTSESG---RYYLGKLTIMWLVKVAKALKVSLMAAATTAALYFPGG-----ATVSA 273

Qy 232 IGYA-----SHDTERLYQETLTFLCYTA 257
: : | : | | | | | |
Db 274 VYMAIARVDVSPVPRKRI-KKATITLPLEENNA 304

RESULT 2
US-10-506-454-493
; Sequence 493, Application US/10506454
; Publication No. US20060068386A1

Query Match	6.6%	Score	88.5	DB	6	Length	263
Best Local Similarity	21.1%	Pred. NO.	1.7				
Matches	50	Conservative	38	Mismatches	84	Indels	65
						Gaps	10

QY	187	HGYPRFRENLR-----	---LYDGDIIWAPADGAFVLT----	TRGGDEDL	224
DB	185	TRKRIQRIILNAGFGRRVIRIERMGL	ENEVVFDEGVDAFGYIAKIGYLRGGSEVL		241

```

RESULT 3
US-11-188-298-10005
; Sequence 10005, Application US/11188298
; Publication No. US20060075522A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: GENES AND USES FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-2115342D,B
; CURRENT APPLICATION NUMBER: US/11/188,298
; PRIORITY FILING DATE: 2005-07-22
; PRIOR APPLICATION NUMBER: 60/592,978
; PRIOR FILING DATE: 2004-07-31
; NUMBER OF SEQ ID NOS: 22569
; SEQ ID NO 10005
; LENGTH: 299

```

```

; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum ATCC 13032
US-11-188-298-10005

```

Query Match	6.6%	Score 88.5	DB 7	Length 299
Best Local Similarity	23.6%	Pred. NO. 2.1		
Matches 60	Conservative 37	Mismatches 96	Indels 61	Gaps 14

```
Qy      |          |          |          |          |          |          |          |          |          |
76 VLRVFTLSRNEIDVERGSKDSWPVEKAA--KKLAFVDRTIFESISYMAALREGIR 1322
```

193	EHINRLVDGI	-----IWAPAID-----	GAFLVLTTR	-----GGDFDLQ	225
-----	------------	-------------------	-----------	--------------	-----

227 DAIDLLDEGHPVH 240

! Sequence 18131, Application US/11188298
! Publication No. US20060075522A1
CUMULATIVE INFORMATION.

```

1  APPLICANT: Aadao, Mark S. et al.
2  TITLE OF INVENTION: GENES AND USES FOR PLANT IMPROVEMENT
3  FILE REFERENCE: 38-21(53452) B
4  CURRENT APPLICATION NUMBER: US/11/186,298
5  CURRENT FILING DATE: 2005-07-22
6  PRIOR APPLICATION NUMBER: 60/592,978
7  PRIOR FILING DATE: 2004-07-31
8  NUMBER OF SEQ. ID NOS: 22569
9  SEQ ID NO 18131
10
11  LENGTH: 315
12
13  TYPE: PRT
14
15  ORGANISM: Corynebacterium glutamicum ATCC 13032
16
17  US-11-186-298-18131

```

Query Match	6.6%	Score	88.5;	DB	7;	Length	315;
Best local similarity	23.6%;	Pred.	No. 2.2;				
Matches	60;	Conservative	37;	Mismatches	96;	Indels	61;
						Gaps	14

QY 76 VRLRVFTLSRNEIDVERGSKDSBPVKEAA---KKLAFVEDRTIFEGYSAAISIEGIR 1322

DB	130	D0DEDVAVMLPBDLVLPIGVMERKQVCR-ABFGG-SVLCAVE---VSEADVSRKIGF	102
QY	193	EHLNRLVLDGDI-----IAPPAID-----GAFVLTTR-----GGNFDLQ	225

Db 243 DAIDLLIDEGHPVH 256

```
RESULT 5
US-11-201-916-18
; Sequence 18, Application US/11201916
; Publication No. US2006003922A1
; GENERAL INFORMATION:
; APPLICANT: Mizzzen, Lee
; APPLICANT: Mizenewski, Jan
; TITLE OF INVENTION: STREPTOCOCCAL HEAT SHOCK PROTEINS OF THE
; TITLE OF INVENTION: HSP60 FAMILY
; NUMBER OF SEQUENCES: 91
; CORRESPONDENCE ADDRESS:
; ADDRESSER: SEED and BERRY LLP
; STREET: 701 Fifth Avenue, 6300 Columbia Center
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/11/201,916
; FILING DATE: 11-AUG-2005
; CLASSIFICATION:
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US/09/001,737
; FILING DATE: 31-DEC-1997
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Sharkey, Richard G.
; REGISTRATION NUMBER: 32,629
; REFERENCE/DOCKET NUMBER: 870109.408
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 539 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
US-11-201-916-18

Query Match      6.4%; Score 88.5; DB 7; Length 539;
Best Local Similarity 25.2%; Pred. No. 4.8;
Matches 66; Conservative 20; Mismatches 77; Indels 99; Gaps 13;

QY 7 DLAPVTEA-----AMAEIELEARTPKRHIAGRVVDVSDPGPYTAAVSTGRLLID 57
DB 289 DLAVTGGGVVNDAGNVLEVEGLFVGSARRVVSODPIYVDGGTAARVANRAK--- 345
QY 58 VKAPTNGVIAHLRASKELVRLRVFTLSRNEIDVERGSKDSW--EPVTEAAKCLA--- 112
DB 346 -----HLRA-----EID-----KSDSDPDRKLAGRLAKLAGV 374
QY 113 -----FYEDRT-----IFEGYSAASIRGIRSSSNPALTLPE 144
DB 375 AVIKGAATETALKEKRESVEDAAVAAKAAVEEGIVGGASLIHQARKLITELRASLTG 434
QY 145 DPRPIPVISQALSE-----LRLAGVDGPFYVLLADVTYKVSF-TSDHGYPIREHLNRL 198
DB 435 DRYLGVDFSEBALAAPFWIAANAAGLDG-----SYVVKVKSLELPAGHGL---NVNLT 483
QY 199 VDGDIWAPAIIDGAF--VLFTTR 218
DB 484 SYGDL-----AADGVIDPVKVT 501

RESULT 6
US-10-506-454-1639
; Sequence 1639, Application US/10506454
```

```
; Publication No. US20060068386A1
; GENERAL INFORMATION:
; APPLICANT: Slesarev, Alexi I
; APPLICANT: Mezhevaya, Katja V
; APPLICANT: Polushin, Nikolai N
; APPLICANT: Shcherbina, Olga V
; APPLICANT: Shakhova, Vera V
; APPLICANT: Malykh, Andrei G
; TITLE OF INVENTION: The Complete Genome and Protein Sequences of the Hyperthermophil
; TITLE OF INVENTION: Methanopyrus kandleri AV19 and Monophyly of Archaeal Methanogens
; FILE REFERENCE: FTD001
; CURRENT APPLICATION NUMBER: US/10/506,454
; CURRENT FILING DATE: 2004-08-31
; PRIOR APPLICATION NUMBER: PCT/US03/06664
; PRIOR FILING DATE: 2003-03-04
; PRIOR APPLICATION NUMBER: 60/361,742
; PRIOR FILING DATE: 2002-03-04
; NUMBER OF SEQ ID NOS: 1722
; SOFTWARE: Patent in Version 3.2
; SEQ ID NO 1639
; LENGTH: 293
; TYPE: PRT
; ORGANISM: Methanopyrus kandleri
US-10-506-454-1639
```

```
Query Match      6.5%; Score 87; DB 6; Length 293;
Best Local Similarity 24.5%; Pred. No. 2.8;
Matches 60; Conservative 30; Mismatches 87; Indels 68; Gaps 14;

QY 9 APVTEAAAEI--ELEAARTFKRHIAGRVVDVSDPGPYTAAVSTGRLLIDVAPNTGVI 66
DB 38 AELTLGLLELGYEVVETTLLEGSLPGRRIELADE---VKGVVGRVYBELLDSDRV- 92
QY 67 ARLPASKPLVRLRVFTLSRNEIDVERGSKDSW-----EPVK-----EAAKCLA 112
DB 93 HHLRLSARCLGAEVKIRYS-----SDWKVEIGYEVDPKHLKRNALVLTLE 139
QY 113 FYEDRTI-----PEGYSAASIRGIRSSSNPALTLPEPREIDV----- 152
DB 140 FVSTREPSTDLVDVIGAAPGWSFPAQMAENV--VAADPA-RLEDRVRELEVVHHLRI 196
QY 153 --ISQALSELRLAGVGPYVLLADVTYKVSFSDHGYPIREHLNRLVVDGIIM----- 205
DB 197 TAHEFVLPENIGVGFGEKXTLLS-DVYS--GNPBDLVNALLERL-ENDVVMGKGVK 252
QY 206 APAID 210
DB 253 APAED 257
```

```
RESULT 7
US-11-087-099-8172
; Sequence 8172, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 8172
; LENGTH: 581
; TYPE: PRT
; ORGANISM: Bradyrhizobium japonicum USDA 110
US-11-087-099-8172

Query Match      6.4%; Score 86.5; DB 7; Length 581;
Best Local Similarity 28.1%; Pred. No. 8.2;
Matches 57; Conservative 30; Mismatches 77; Indels 39; Gaps 13;
```



```

; SEQ ID NO 8396
; LENGTH: 538
; TYPE: PRT
; ORGANISM: Sphingomonas paucimobils
US-11-188-298-8396

Query Match          6.2%; Score 84; DB 7; Length 538;
Best Local Similarity 20.5%; Pred. No. 12;
Matches 70; Conservative 33; Mismatches 104; Indels 134; Gaps 14;

QY      13 EAANAA-IELEAAATPKKHINGRRVVVSDGPGVTAAVSTG-----RLIDVKAPTVGV 65
Db       143 EPAAAEVPLALIRAY--HNA-----MQARGVFVSVPLDDMDRACTPIDVSVSQDV 194
QY      66 IAHL----RASKPLVRLRVPTLSRNEIDVERGSKSDWEPEAKKLAFAV----- 114
Db       195 MASPLALGAAAEALDKKRSFVTVAG---AGVGRNARALAKLEAHKARVFAVAMARA 251
QY      115 ---EDRIIFEGYSAASTIEGIRSAASN----- 137
Db       252 VFPEHDPLFAGFLPASREAAVSAALKGHDVLVLGAPVFTYHVEGRGPVPEGASLIQIVD 311
QY      138 -----PALTLPEDPREIIPD-----VISQALSE 159
Db       312 DPSVAATPVGTAIGVADALARLSAITSAPFRAVFDALPAPNVGSGTAISEALLM 371
QY      160 LRLAGVDGPYSVLISADVYTKVSETSDHGYPIREHLNRLVDGDIIMAPALDGAFLVLTTRG 219
Db       372 QQLATL-RPKGIITISEAPSTRGPMHDH-PIIR-----LDEDFWATASG 413
QY      220 GDEPLQLGTVVAIGYASHDDTTERLYLQETITLICTAAAS 260
Db       414 G--LGFALPAVAGAAIAQDPD-----TVLCLLDGS 442

RESULT 13
US-11-079-463-6012
; Sequence 6012, Application US/11079463
; Publication No. US20060073161A1
; GENERAL INFORMATION:
; APPLICANT: Gary L. Bretton
; TITLE OR INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO BACTERIOIDS FROM
; FILE REFERENCE: PATH00-03DIV2
; CURRENT APPLICATION NUMBER: US/11/079,463
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/128,705
; PRIOR FILING DATE: 1999-04-09
; PRIOR APPLICATION NUMBER: US 09/540,209
; PRIOR FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 10444
; SEQ ID NO 6012
; LENGTH: 901
; TYPE: PRT
; ORGANISM: B. fragilis
US-11-079-463-6012

Query Match          6.2%; Score 84; DB 7; Length 901;
Best Local Similarity 23.0%; Pred. No. 26;
Matches 59; Conservative 37; Mismatches 105; Indels 56; Gaps 11;

QY      31 IAGRRVVDSDPGGPVTAAVSTG-----RLIDVKAPTVGV----- 66
Db       48 IAGTTGTAVDVGDNATLTLSNGTITLEVKIKIGYKTLRMEVKKKAAATLTFELEVDAT 107
QY      67 -----AHLRASKPLVRLRVPTLSRNEIDVERGSKSDWE---PYEAAKKLAFAV 114
Db       108 LDATVAVARNLGEKRLDGRQKATLA---IENM--GAKEMTLKGISNVQDGVKCTNGI 162
QY      115 E-----DRIIFEGYSAASTIEGIRSAASNPALTL-PEDPREIFDVISQALSELRLAGV 165
Db       163 SIASAGQLIVRGIDRISTYTLNGLPIASFPNPDKLPLDLFPASTKNTITVSKVYAGA 222

```

QY 166 DGPYVLLSADVTYKVSFSDH---GYPIREHLNRLVDGDIWAPALDGAFLV--TRGGD 221
DB 223 FADYSG-AHIDISTKENTGSDPFSIGFNVGGRFN--TVGKDPYYSDRKGILFSGNLANKD 280
QY 222 FDLQLGTDAVIGYASHD 238
DB 281 RILAMGKSEPRDYARNN 297

RESULT 14

US-10-453-372-1142
/ Sequence 1142, Application US/10453372
/ Publication No. US20060003323A1
/ GENERAL INFORMATION:
/ APPLICANT: Alcobrook, et al.
/ TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
/ FILE REFERENCE: 21402-589 A
/ CURRENT APPLICATION NUMBER: US/10/453,372
/ PRIOR FILING DATE: 2003-06-03
/ PRIOR APPLICATION NUMBER: 09/789390
/ PRIOR FILING DATE: 2001-02-23
/ PRIOR APPLICATION NUMBER: 60/185967
/ PRIOR FILING DATE: 2000-03-01
/ PRIOR APPLICATION NUMBER: 09/823187
/ PRIOR FILING DATE: 2001-03-29
/ PRIOR APPLICATION NUMBER: 60/195792
/ PRIOR FILING DATE: 2000-03-10
/ PRIOR APPLICATION NUMBER: 09/839446
/ PRIOR FILING DATE: 2001-03-19
/ PRIOR APPLICATION NUMBER: 60/199476
/ PRIOR FILING DATE: 2000-03-25
/ PRIOR APPLICATION NUMBER: 09/863776
/ PRIOR FILING DATE: 2001-05-23
/ PRIOR APPLICATION NUMBER: 60/208263
/ PRIOR FILING DATE: 2000-05-31
/ PRIOR APPLICATION NUMBER: 09/939398
/ PRIOR FILING DATE: 2001-08-24
/ PRIOR APPLICATION NUMBER: 60/227800
/ PRIOR FILING DATE: 2000-08-25
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 1609
/ SOFTWARE: Curaseq1et version 0.1
/ SEQ ID NO 1142
/ LENGTH: 4913
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-453-372-1142

Query Match 6.2%; Score 84; DB 6; Length 4913;
Best Local Similarity 21.4%; Pred. No. 2.9e+02;

Matches 66; Conservative 50; Mismatches 106; Indels 86; Gaps 17;

QY 10 PYTEAMAEIELEAARFKRHINAGRRVDV---SDPG--GPTAAVSTGRLLDVAKPTN 63
DB 2052 PVEAQLLYKVEIN-----ENTLGTDTDIQVFAADGDEGTNGQVYGVNGN--TNOEFRID 2105
QY 64 GVIAHLRASKEPLVRLVP--FTLSRNEIDDERGSKD-----SDWEPYKEAA 108
DB 2106 SVTGATTVAKPLDREKTPYHLT---VQATDRGSTRTDTSTVSYVLNDINDVPVPELS 2162
QY 109 KKLAFVED-----RTIFEGYSAASIRGIRASASNP--LTLPEDPREIPDVI----- 153
DB 2163 PYSVNPENLGTLPRTILGTASPC---VRFASASKAYFTTIPEDAPGTGTVLNVASDA 2218
QY 154 ---SQAISERLAGVDGPGS-----VLASA-----DYTKVSTSDHG--YPIREH 194
DB 2219 DASKNAVISYRIIGNSQFINPSTGOITTSALDRETKDNYTLVVVCSDGSPPEPLSSS 2278
QY 195 LNLVD--GDI-----IWAPALDGAFLVTRGGDFDLQLGTDAVIGYASHD 238
DB 2279 TSVLVTVTTVDHNDPFRQHHPIVYTHIPSPILPSGFVAATVTDAD--IGNSSLHYSLSG 2336
QY 239 TDTERLVL 246

DB 2337 RNSEKPHI 2344

RESULT 15

US-10-453-372-1132
/ Sequence 1132, Application US/10453372
/ Publication No. US20060003323A1
/ GENERAL INFORMATION:
/ APPLICANT: Alcobrook, et al.
/ TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
/ FILE REFERENCE: 21402-589 A
/ CURRENT APPLICATION NUMBER: US/10/453,372
/ PRIOR FILING DATE: 2003-06-03
/ PRIOR APPLICATION NUMBER: 09/789390
/ PRIOR FILING DATE: 2001-02-23
/ PRIOR APPLICATION NUMBER: 60/185967
/ PRIOR FILING DATE: 2000-03-01
/ PRIOR APPLICATION NUMBER: 09/823187
/ PRIOR FILING DATE: 2001-03-29
/ PRIOR APPLICATION NUMBER: 60/195792
/ PRIOR FILING DATE: 2000-03-10
/ PRIOR APPLICATION NUMBER: 09/839446
/ PRIOR FILING DATE: 2001-03-19
/ PRIOR APPLICATION NUMBER: 60/199476
/ PRIOR FILING DATE: 2000-03-25
/ PRIOR APPLICATION NUMBER: 09/863776
/ PRIOR FILING DATE: 2001-05-23
/ PRIOR APPLICATION NUMBER: 60/208263
/ PRIOR FILING DATE: 2000-05-31
/ PRIOR APPLICATION NUMBER: 09/939398
/ PRIOR FILING DATE: 2001-08-24
/ PRIOR APPLICATION NUMBER: 60/227800
/ PRIOR FILING DATE: 2000-08-25
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 1609
/ SOFTWARE: Curaseq1et version 0.1
/ SEQ ID NO 1132
/ LENGTH: 4961
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-453-372-1132

Query Match 6.2%; Score 84; DB 6; Length 4961;
Best Local Similarity 21.4%; Pred. No. 2.9e+02;

Matches 66; Conservative 50; Mismatches 106; Indels 86; Gaps 17;

QY 10 PYTEAMAEIELEAARFKRHINAGRRVDV---SDPG--GPTAAVSTGRLLDVAKPTN 63
DB 2100 PVEAQLLYKVEIN-----ENTLGTDTDIQVFAADGDEGTNGQVYGVNGN--TNOEFRID 2153
QY 64 GVIAHLRASKEPLVRLVP--FTLSRNEIDDERGSKD-----SDWEPYKEAA 108
DB 2154 SVTGATTVAKPLDREKTPYHLT---VQATDRGSTRTDTSTVSYVLNDINDVPVPELS 2210
QY 109 KKLAFVED-----RTIFEGYSAASIRGIRASASNP--LTLPEDPREIPDVI----- 153
DB 2211 PYSVNPENLGTLPRTILGTASPC---VRFASASKAYFTTIPEDAPGTGTVLNVASDA 2266
QY 154 ---SQAISERLAGVDGPGS-----VLASA-----DYTKVSTSDHG--YPIREH 194
DB 2267 DASKNAVISYRIIGNSQFINPSTGOITTSALDRETKDNYTLVVVCSDGSPPEPLSSS 2326
QY 195 LNLVD--GDI-----IWAPALDGAFLVTRGGDFDLQLGTDAVIGYASHD 238
DB 2327 TSVLVTVTTVDHNDPFRQHHPIVYTHIPSPILPSGFVAATVTDAD--IGNSSLHYSLSG 2384
QY 239 TDTERLVL 246
DB 2385 RNSEKPHI 2392

Search completed: April 14, 2006, 18:42:09

Job time : 13.3554 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocelebration Ltd.

OM protein - protein search, using sw model

Run on: April 14, 2006, 17:03:40 ; Search time 115.851 Seconds
(without alignments)
1133.996 Million cell updates/sec

Title: US-10-620-246-42
Perfect score: 1621
Sequence: 1 MKRSLALRLALMTALSLFGL.....WGSWAPQLGAMSGDIVGAIR 299

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_21.*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*
9: geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1621	100.0	299	2	AAW72899 Mycobacte
2	1621	100.0	299	2	AAW72899 Mycobacte
3	1621	100.0	299	8	ADRI2610 Gene vacc
4	1616	99.7	299	2	AAW18162 Mycobacte
5	1616	99.7	299	2	AAW63037 Mycobacte
6	1616	99.7	299	7	AAE39354 M. tuberc
7	1616	99.7	299	8	ADRI2612 Gene vacc
8	1616	99.7	299	8	ADU64327 Mycobacte
9	1462	90.2	266	7	ADM40808 Mycobacte
10	1445	89.1	266	9	ADZ97846 Mycobacte
11	1394	86.0	256	7	ADFI1940 Mycobacte
12	564.5	34.8	340	4	AAAB47557 Ag85C. 12
13	564.5	34.8	340	5	AAAM50731 Mycobacte
14	561.5	34.6	340	2	AAAR26167 85-C. 3/2
15	561.5	34.6	340	2	AAAY14857 Antigen 8
16	561.5	34.6	340	5	ABW73463 M. tubercu
17	548	33.8	333	3	AAV14856 Antigen 8
18	548	33.8	333	5	ABW73462 M. leprae
19	532	32.8	337	6	ABU36126 Protein e
20	532	32.8	337	6	ABU56407 Mycobacte
21	532	32.8	338	2	AAW18165 Mycobacte
22	532	32.8	338	2	AAW63033 Mycobacte
23	532	32.8	338	2	AAV14852 Antigen 8
24	532	32.8	338	2	AAV14854 Antigen 8

25	532	32.8	338	4	AAW82789 Mycobacte
26	532	32.8	338	4	AAW47555 Ag85A. 12
27	532	32.8	338	5	AAW50729 Mycobacte
28	532	32.8	338	5	ABW73458 M. tubercu
29	532	32.8	338	5	ABW73460 M. bovis 8
30	532	32.8	338	7	AAE39295 M. tuberc
31	532	32.8	338	7	ADP45194 M. tuberc
32	532	32.8	338	8	ADO36879 M. tuberc
33	532	32.8	338	8	ADRI2616 Gene vacc
34	532	32.8	338	8	ADRI2618 Gene vacc
35	532	32.8	338	8	ADU64323 Mycobacte
36	529.5	32.7	766	9	ABW26442 M. tuberc
37	527.5	32.5	674	9	ABW26432 M. tuberc
38	526	32.4	338	2	AAW11295 Recombina
39	526	32.4	353	2	AAW11295 Recombina
40	523.5	32.3	487	9	ABW26444 M. tuberc
41	523	32.3	330	2	AAV14850 Antigen 8
42	523	32.3	330	5	ABW73456 M. leprae
43	523	32.3	496	9	ABW26448 M. tuberc
44	523	32.3	782	9	ABW26436 M. tuberc
45	520	32.1	325	8	ADRI2620 Gene vacc

ALIGNMENTS

RESULT 1	
AAW72899	standard; protein; 299 AA.
ID	AAW72899
AC	AAW72899;
XX	
DT	21-JAN-1999 (first entry)
XX	
DE	Mycobacterium tuberculosis antigen MPT51.
XX	
KW	Mycobacterium tuberculosis; antigen; vaccine; immunological; immunogen; infection.
XX	
OS	Mycobacterium tuberculosis.
XX	
FN	W09844119-A1.
XX	
PD	08-OCT-1998.
XX	
PF	01-APR-1998; 98WC-DK000132.
XX	
PR	02-APR-1997; 97DK-00000376.
PR	18-APR-1997; 97DK-0044624P.
PR	10-NOV-1997; 97DK-00001277.
PR	05-JAN-1998; 98US-0070488P.
XX	
PA	(STAT-) STATENS SERUM INST.
XX	
PI	Andersen P, Nielsen R, Rosenkrands I, Wellingh K, Rasmussen PB, Oettinger T, Florio W,
XX	WPI: 1998-542705/46.
DR	N-PSDB; AAW63923.
XX	
PT	New isolated mycobacteria polypeptides and nucleic acids - used for developing products for the diagnosis of or vaccination against mycobacterial infections, particularly tuberculosis.
PT	mycobacterial infections, particularly tuberculosis.
XX	
PS	Claim 1, Page 151-152; 163pp; English.
XX	
CC	The present sequence represents a Mycobacterium tuberculosis protein. Products from the present invention, which describes protein fragments and nucleic acid fragments derived from M. tuberculosis, can be used in the detection of and prevention of mycobacterial infections. In particular, the proteins and nucleic acids can be used for the diagnosis of or vaccination against tuberculosis caused by M. tuberculosis, M. africanum or M. bovis
CC	

XX Sequence 299 AA;
 SQ Query Match 100.0%; Score 1621; DB 2; Length 299;
 Best Local Similarity 100.0%; Pred. No. 7.2e-121;
 Matches 299; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKGRSALLRALMTIALSFGIGVAVAAEPTAKAPYENLMVPSMGRDIPVAFLAGCPH 60
 |||||
 DB 1 MKRSALLRLMTIALSFGIGVAVAAEPTAKAPYENLMVPSMGRDIPVAFLAGCPH 60
 |||||

QY 61 AVTLDAFNAGPVSVMVTAGNMTLAGGISVAPAGAGYSMTYTWEDDGSKOMDTFL 120
 |||||
 DB 61 AVTLDAFNAGPVSVMVTAGNMTLAGGISVAPAGAGYSMTYTWEDDGSKOMDTFL 120
 |||||

QY 121 SAELPDMILANRGLAPGHAAGAAGGYGAMALLAHPDRFGPAGSMGFLYPSNTTN 180
 |||||
 DB 121 SAELPDMILANRGLAPGHAAGAAGGYGAMALLAHPDRFGPAGSMGFLYPSNTTN 180
 |||||

QY 181 GATAAGMOQFGVDITNGMGAPOLGKWKHDPVTHASLLAQNNTRVVMSPTNPGASDPA 240
 |||||
 DB 181 GATAAGMOQFGVDITNGMGAPOLGKWKHDPVTHASLLAQNNTRVVMSPTNPGASDPA 240
 |||||

QY 241 AMIGOTAEAMGNSRMFTNYRVSVGHNHFDPPASGDNMGSWAPOLGAMSGDIVGAIR 299
 |||||
 DB 241 AMIGOTAEAMGNSRMFTNYRVSVGHNHFDPPASGDNMGSWAPOLGAMSGDIVGAIR 299
 |||||

RESULT 2
 AA121916
 ID AA121916 standard; protein; 299 AA.
 XX
 AC AA121916;
 XX
 DT 06-SRP-1999 (first entry)
 XX
 DE Amino acid sequence of antigen MPT51.
 XX
 KM Immunogenic; Mycobacterium tuberculosis; immune response; infection;
 KM tuberculosis; fusion polypeptide; T-cell epitope; ESAT-6; MPT59; TB;
 KM pharmaceutical; vaccination; M. africanum; M. bovis; CFP7A; CFP30A;
 KM CFP7B; CFP13; CFP27; CFP30A; RD1-ORF; CFP10A; CFP16; CFP19; CFP23;
 KM CFP25A; CFP30B; CFP7B.
 XX
 OS Mycobacterium tuberculosis.
 XX
 PN MO924577-A1.
 XX
 PD 20-MAY-1999.
 XX
 PF 08-OCT-1998; 98MO-DK000438.
 XX
 PR 10-NOV-1997; 97DK-00001277.
 PR 05-JAN-1998; 98US-0070488P.
 PR 01-APR-1998; 98MO-DK000132.
 XX
 PA (STAT-) STATENS SERUM INSTR.
 XX
 PI Andersen P, Skjot R;
 XX
 DR WPI; 1999-347282/29.
 DR N-PSDB; AAX81026.
 XX
 PT New immunogenic fragment of Mycobacterium tuberculosis.
 XX
 PS Example 3; Page 172-173; 265pp; English.
 XX
 CC The invention describes a substantially pure immunogenic polypeptide
 CC fragment (1) from Mycobacterium tuberculosis that is able to evoke a
 CC protective immune response against infections by mycobacteria belonging
 CC to the tuberculosis complex. The invention provides a (1) fusion
 CC polypeptide comprising at least one polypeptide fragment (1) and at least
 CC one fusion partner; (2) a fusion polypeptide fragment comprising a T-cell

CC epitope from M. tuberculosis protein ESAT-6, or MPT59 and a second
 CC different amino acid sequence from M. tuberculosis, and/or including a
 CC sequence which protects the first amino acid sequence from in vivo
 CC degradation or post-translational processing; (3) a nucleic acid fragment
 CC that encodes the above polypeptides. The polypeptides and nucleic acid
 CC are useful as pharmaceuticals, for diagnosis of and as antigens for
 CC vaccination against TB caused by Mycobacterium tuberculosis, africanum or
 CC bovis. The polypeptides are also useful for diagnosing ongoing or
 CC previous sensitization in an animal with bacteria belonging to the
 CC tuberculosis complex. The invention also describes the use of CFP7A or
 CC CFP30A or a T-cell epitope of for the induction of a strong immune
 CC response in a mammal; use of CFP7B, CFP19 or MPT59-ESAT6 or a T-cell
 CC epitope of for diagnosis of TB in a mammal by performing a DTH type skin
 CC test; use of CFP27, CFP30A, RD1-ORF2, RD1-ORF3, RD1-ORF5, MPT59-ESAT6,
 CC ESAT6-MPT59, CFP10A, CFP16, CFP19, CFP23, CFP25A, CFP30B, CFP7B or a T-
 CC cell epitope of for the preparation of an immunological composition; and
 CC for the preparation of a subunit vaccine

XX Sequence 299 AA;
 SQ Query Match 100.0%; Score 1621; DB 2; Length 299;
 Best Local Similarity 100.0%; Pred. No. 7.2e-121;
 Matches 299; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKGRSALLRALMTIALSFGIGVAVAAEPTAKAPYENLMVPSMGRDIPVAFLAGCPH 60
 |||||
 DB 1 MKRSALLRLMTIALSFGIGVAVAAEPTAKAPYENLMVPSMGRDIPVAFLAGCPH 60
 |||||

QY 61 AVTLDAFNAGPVSVMVTAGNMTLAGGISVAPAGAGYSMTYTWEDDGSKOMDTFL 120
 |||||
 DB 61 AVTLDAFNAGPVSVMVTAGNMTLAGGISVAPAGAGYSMTYTWEDDGSKOMDTFL 120
 |||||

QY 121 SAELPDMILANRGLAPGHAAGAAGGYGAMALLAHPDRFGPAGSMGFLYPSNTTN 180
 |||||
 DB 121 SAELPDMILANRGLAPGHAAGAAGGYGAMALLAHPDRFGPAGSMGFLYPSNTTN 180
 |||||

QY 181 GATAAGMOQFGVDITNGMGAPOLGKWKHDPVTHASLLAQNNTRVVMSPTNPGASDPA 240
 |||||
 DB 181 GATAAGMOQFGVDITNGMGAPOLGKWKHDPVTHASLLAQNNTRVVMSPTNPGASDPA 240
 |||||

QY 241 AMIGOTAEAMGNSRMFTNYRVSVGHNHFDPPASGDNMGSWAPOLGAMSGDIVGAIR 299
 |||||
 DB 241 AMIGOTAEAMGNSRMFTNYRVSVGHNHFDPPASGDNMGSWAPOLGAMSGDIVGAIR 299
 |||||

RESULT 3
 ADRI2610
 ID ADRI2610 standard; protein; 299 AA.
 XX
 AC ADRI2610;
 XX
 DT 04-NOV-2004 (first entry)
 XX
 DE Gene vaccine protein #34.
 XX
 KM antitubercular; tuberculosis; anti-HIV; antibacterial;
 KM haemostatic; protozoacide; antiinflammatory; neuroprotective; vitruclide;
 KM gene vaccine; ubiquitin; antigen; intracellular parasite;
 KM protozoan parasite infection; malaria; tuberculosis; toxoplasmosis;
 KM trypanosomiasis; AIDS; cytomegalovirus infection; achlamydia disease;
 KM Rickettsia; leishmaniasis; Ebola hemorrhagic fever; Trypanosoma;
 KM Chagas disease; Japanese encephalitis; influenza; rubella; dengue virus;
 KM poliomyelitis; Herpes virus; severe acute respiratory syndrome.
 XX
 OS Unidentified.
 XX
 PN MO2004067040-A1.
 XX
 PD 12-AUG-2004.
 XX
 PF 30-JAN-2004; 2004MO-JP000975.
 XX
 PR 31-JAN-2003; 2003JP-00023507.


```

AAW63037
ID AAW63037 standard; protein; 299 AA.
XX
AC AAW63037;
XX
DT 23-OCT-1998 (first entry)
XX
DE Mycobacterium tuberculosis 24 kd protein sequence.
XX
KM Mycobacterium tuberculosis; vaccination; extracellular product;
KW immunodominant epitope; interleukin-12; M59; immune response;
XX opsonising humoral response; intracellular pathogen.
XX
OS Mycobacterium tuberculosis.
XX
PN W09831388-A1.
XX
PD 23-JUL-1998.
XX
PF 15-JAN-1998; 98WO-US000942.
XX
PR 21-JAN-1997; 97US-00786533.
XX
PA (REGC ) UNIV CALIFORNIA.
XX
PI Horwitz MA, Harth G, Lee B;
XX
DR WPI; 1998-413815/35.
XX
DR N-PSDB; AAW42600.
XX
PT Vaccines against Mycobacterium containing major extracellular proteins -
PT used to, e.g. induce protective and therapeutic immune responses, and for
XX detecting an immune response.
XX
PS Example 2; Page 45-46; 236pp; English.
XX
CC This represents a Mycobacterium tuberculosis 30 kd protein. The invention
CC provides an agent for vaccinating mammals against Mycobacterium. The
CC agent comprises at least one of the major abundant extracellular 110, 80,
CC 71, 58, 45, 32A, 32B, 30, 24, 23.5, 23, 16, 14 or 12 kd proteins of M.
CC tuberculosis, or at least 1 of their immunodominant epitopes and
CC interleukin-12 (IL-12) or M59 as adjuvants. The agent containing the
CC nucleic acid encoding the extracellular products are used to raise a
CC protective or therapeutic immune response against Mycobacterium,
CC specifically M. tuberculosis. The immunodominant epitopes can also be
CC used (typically in a cutaneous hypersensitivity test) to detect an immune
CC response to vaccination. Preparation of the agent does not require
CC selection of the most immunogenic products, so large scale production and
CC purification are easy, resulting in a consistent, standardised
CC formulation, having lower toxicity than killed or attenuated vaccines.
CC The agents provide a rapid and effective response (including a strong
CC cell-mediated component) and are safe even in immunocompromised subjects.
CC They prevent development of an opsonising humoral response that might
CC spread intracellular pathogens
XX
SQ Sequence 299 AA;

Query Match          99.7%; Score 1616; DB 2; Length 299;
Best Local Similarity 99.7%; Pred. No. 1.8e-120;
Matches 298; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 181 GATAGAGCGCTDTGKMGAPOLGRKWKHDPVWVHSLAQNNTRYVWWSPTNPGASDPA 240
DB 181 GATAGAGCGCTDTGKMGAPOLGRKWKHDPVWVHSLAQNNTRYVWWSPTNPGASDPA 240
QY 241 AMTGTABAMGNSRMRYNQRYSVGHNGHPDPFASDNGMGWAPOLGAMSGDIYGAIR 299
DB 241 AMTGTABAMGNSRMRYNQRYSVGHNGHPDPFASDNGMGWAPOLGAMSGDIYGAIR 299

RESULT 6
AAE39354
ID AAE39354 standard; protein; 299 AA.
XX
AC AAE39354;
XX
DT 18-DEC-2003 (first entry)
XX
DE M. tuberculosis extracellular 24 kd protein.
XX
KM Vaccine; antibacterial; fungicide; protozoacide; immunostimulant;
KW virucide; therapy.
XX
OS Mycobacterium tuberculosis.
XX
PN US6599510-B1.
XX
PD 29-JUL-2003.
XX
PF 21-SEP-1998; 98US-00157689.
XX
PR 23-NOV-1993; 93US-00156358.
PR 12-AUG-1994; 94US-00289667.
PR 23-MAY-1995; 95US-00447398.
PR 31-OCT-1995; 95US-00551149.
PR 06-DEC-1995; 95US-00566357.
PR 23-MAY-1996; 96US-00652842.
XX
PA (REGC ) UNIV CALIFORNIA.
XX
PI Horwitz MA, Harth G;
XX
DR WPI; 2003-669607/63.
DR N-PSDB; AAD59705.
XX
PT A nucleic acid encoding an abundant extracellular protein of
PT Mycobacterium tuberculosis useful as vaccines for generating protective
PT or therapeutic immune response against viral, bacterial, fungal or
PT protozoal infections.
XX
PS Example 2; Col 33-36; 82pp; English.
XX
CC The invention relates to a novel nucleic acid encoding an abundant
CC extracellular protein of Mycobacterium tuberculosis useful as vaccines
CC for generating protective or therapeutic immune response against viral,
CC bacterial, fungal and protozoal infections. They are also used as
CC immunotherapeutic agents. The present sequence is M. tuberculosis strain
CC Bkman extracellular protein
XX
SQ Sequence 299 AA;

Query Match          99.7%; Score 1616; DB 7; Length 299;
Best Local Similarity 99.7%; Pred. No. 1.8e-120;
Matches 298; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

Db 121 SAEIPDWLAANRGLAPGSHAIVGAAGGYGAMALAAFPDRFGPAGSMGFLYPSNTTTN 180
 Qy 181 GATAAGNQOFGVDYTNMGWGAPOLGKMKHDPWVHASLLAQNNTTRVWVMSPTNGASDPA 240
 Db 181 GATAAGNQOFGVDYTNMGWGAPOLGKMKHDPWVHASLLAQNNTTRVWVMSPTNGASDPA 240
 Qy 241 AMIGQTAEMGNSMFPNQRYSVGGHNGHDPFPASGDNMGSWAPOLGAMSGDIVGAIR 299
 Db 241 AMIGQTAEMGNSMFPNQRYSVGGHNGHDPFPASGDNMGSWAPOLGAMSGDIVGAIR 299

RESULT 7
 ADR12612 ID ADR12612 standard; protein; 299 AA.

AC ADR12612;
 XX
 DT 04-NOV-2004 (first entry)
 XX
 XX Gene vaccine protein #35.

KW antimarial; antitubercular; tuberculostatic; anti-HIV; antibacterial;
 KW haemostatic; protozoacide; antiinflammatory; neuroprotective; virucide;
 KW gene vaccine; ubiquitin; antigen; intracellular parasite;
 KW protozoan parasite infection; malaria; tuberculosis; toxoplasmosis;
 KW trypanosomiasis; AIDS; cytomegalovirus infection; achiyandia disease;
 KW Rickettsia; leishmaniasis; Ebola hemorrhagic fever; Trypanosoma;
 KW Chagas disease; Japanese encephalitis; Influenza; rubella; dengue virus;
 KW poliomyelitis; Herpes virus; severe acute respiratory syndrome.

XX Unidentified.

XX W02004067040-A1.

XX 12-AUG-2004.

XX 30-JAN-2004; 2004MO-JP000975.

XX 31-JAN-2003; 2003JP-00023507.

XX (KYUS-) KYUSHU TLO CO LTD.

XX Himeno K, Ishii K;

XX WPI; 2004-594036/57.

XX N-PSDB; ADR12611.

XX Gene vaccine having nucleic acid encoding ubiquitin and antigen protein
 PT of intracellular parasite, useful for treating parasite infections such
 as malaria, tuberculosis, toxoplasmosis.

XX Disclosure; SEQ ID NO 79; 266pp; Japanese.

XX A gene vaccine (I) comprises nucleic acid sequence encoding ubiquitin,
 CC and a nucleic acid sequence encoding the antigen protein of an
 CC intracellular parasite containing T-cell target sequence. (I) is useful
 CC for preventing or treating the disease resulting from intracellular
 CC protozoan parasite infection such as malaria, tuberculosis,
 CC toxoplasmosis, trypanosomiasis, AIDS, cytomegalovirus infection,
 CC achlamydia disease, infections caused by Rickettsia, leishmaniasis,
 CC Ebola hemorrhagic fever, Trypanosoma infections, Chagas disease, Japanese
 CC encephalitis, influenza, rubella and dengue viral infections,
 CC poliomyelitis, Herpes virus (alpha) infections, or severe acute
 CC respiratory syndrome. This sequence represent a protein used in the
 CC method of the invention.

XX Sequence 299 AA;

Query Match 99.7%; Score 1616; DB 8; Length 299;
 Best Local Similarity 99.7%; Pred. No. 1.8e-120;
 Matches 299; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MKGRSALLRALMTAALSFGLGVAVAAEPTAKAAYENLWVSPSGMRDIPVAFLAGGP 60
 Db 1 MKGRSALLRALMTAALSFGLGVAVAAEPTAKAAYENLWVSPSGMRDIPVAFLAGGP 60
 Qy 61 AVYLLDAFNAQPDVSNWVTAGNANTTLAKGISVVAAGAGYSNWTTMEDQSGNQWTF 120
 Db 61 AVYLLDAFNAQPDVSNWVTAGNANTTLAKGISVVAAGAGYSNWTTMEDQSGNQWTF 120
 Qy 121 SAEIPDWLAANRGLAPGSHAIVGAAGGYGAMALAAFPDRFGPAGSMGFLYPSNTTTN 180
 Db 121 SAEIPDWLAANRGLAPGSHAIVGAAGGYGAMALAAFPDRFGPAGSMGFLYPSNTTTN 180
 Qy 181 GATAAGNQOFGVDYTNMGWGAPOLGKMKHDPWVHASLLAQNNTTRVWVMSPTNGASDPA 240
 Db 181 GATAAGNQOFGVDYTNMGWGAPOLGKMKHDPWVHASLLAQNNTTRVWVMSPTNGASDPA 240
 Qy 241 AMIGQTAEMGNSMFPNQRYSVGGHNGHDPFPASGDNMGSWAPOLGAMSGDIVGAIR 299
 Db 241 AMIGQTAEMGNSMFPNQRYSVGGHNGHDPFPASGDNMGSWAPOLGAMSGDIVGAIR 299

RESULT 8

ADU64327 ID ADR12612 standard; protein; 299 AA.

ADU64327;

XX 10-FEB-2005 (first entry)

XX Mycobacterium tuberculosis 24 KD protein, SEQ ID 166.

XX Antibacterial; Tuberculostatic; Protozoacide; Fungicide; Virucide;

XX Anti-HIV; Vaccine.

XX Mycobacterium tuberculosis.

XX US2004228873-A1.

XX 18-NOV-2004.

XX 27-OCT-2003; 2003US-00695155.
 XX 23-NOV-1993; 93US-00156358.

XX 12-AUG-1994; 94US-00289667.

XX 23-MAY-1995; 95US-00447398.

XX 20-OCT-1995; 95US-00545926.

XX 31-OCT-1995; 95US-00551149.

XX 23-MAY-1996; 96US-00652842.

XX 06-DEC-1996; 96US-00568357.

XX 21-JAN-1997; 97US-00786533.

XX (HORN/) HORWITZ M A.

XX (HART/) HARTH G.

XX (LEEB/) LEB B.

XX Horwitz MA, Harth G, Lee B;

XX WPI; 2004-813210/80.

XX N-PSDB; ADU64256.

XX New vaccinating agent for promoting an immune response against infectious

XX Mycobacterium pathogen and HIV infection using majorly abundant

XX extracellular products from Mycobacterium tuberculosis.

XX Example 2; SEQ ID NO 166; 109pp; English.

The present invention relates to a vaccinating agent for promoting an
 CC immune response in a mammalian host against an infectious Mycobacterium
 CC pathogen. The vaccinating agent comprises at least a portion or at least
 CC one immunodominant epitope of at least one majorly abundant extracellular
 CC product selected from M. tuberculosis 110, 80, 71, 58, 45, 328, 30,
 CC 24, 23.5, 23, 16, 14 or 12 KD protein and their analogs, homologs, and
 CC subunits, and an IL-12 or MF 59 adjuvant. The compositions of the present

CC invention are useful for stimulating a protective immune response in
CC mammalian hosts against intracellular pathogens, such as bacteria,
CC protozoa, viruses and fungus, particularly *Mycobacterium tuberculosis* and
CC the HIV virus. The present sequence is a *M. tuberculosis* extracellular
CC protein.

CC comprising CD8 T-cells with the peptide in question and detecting whether
CC the CD8 T-cells recognize the peptide, by detecting the expression of a
CC substance by the T cell which indicates that the T cell has recognized
CC the peptide or by detecting lysis by T cells of cells that present the
CC peptide on their surface, the detection of lysis indicating that the T
CC cells have recognized the peptide

XX Sequence 340 AA;

Query Match 34.8%; Score 564.5; DB 4; Length 340;
Best Local Similarity 41.6%; Pred. No. 1.3e-36;
Matches 131; Conservative 38; Mismatches 119; Indels 27; Gaps 8;

QY 4 RSA-----LRALMIA-----LSFGL-----GGVAVAEPTAKAPYENIMVPSPSMGNDIPV 52
DB 11 RSNATTLPRRLAIAMGAVLVYGLVGTGPGPATNAGAFSRGPLEVEYIQVPSASMGNDIKY 70
QY 53 AFLAGPRAVYLLDAFNAGPDVSNWVTAGNANTLAGKGISVVAAPAGAYSMNTNEQDG 112
DB 71 QFGGGGPHAVYLLDGLRAQDDYNGMDINTPAFEYYSQGLSVIMPVGQSSFYTDWYOPS 130

QY 113 SK-----QMDTFLSAELPDMILAANRGLAPGHAIVGAAGGYGAMALAAFPDRRGA 165
DB 131 QSNQNTYTKMETFLTRMPAMLOANKGVSPTGNAVGLSMGSGSALITLAATYPOQFPYA 190

QY 166 GSWGFLYPSNTTNGAIAAGMOQFGVDNCGMKGAPQLGRWKMHPWYHASTLLAQNTR 225
DB 191 ASLSGFLNPSBGWMPFLIGLANNDSGYNANSMWGPSDPAMKRNPMVQIPRLVANTR 250

QY 226 VWWV-----SPTNPGASD-PAAMI-GQTAEAMGNSRMFYNOYRSVGHNGHFDPPASGNDG 279
DB 251 IWYCGNGTFPSDLGGDNIPAKFLLEGTL---LRTNQTFRDYAADGGNGVFNFPNGTHS 307

QY 280 WGSNAPQLGAMSGDI 294
DB 308 WPYMNEQLVAMKADI 322

RESULT 13
AAMS0731
ID AAMS0731 standard; protein; 340 AA.

XX AAMS0731;
XX 18-APR-2002 (first entry)

DE Mycobacterium tuberculosis immunodominant Mtb protein FbpC2 Ag85C.
XX Mycobacterium tuberculosis immunogen; mycobacteria; immunisation; vaccine.
XX OS

XX Mycobacterium tuberculosis.
XX Key
XX Location/Qualifiers

FT Peptide 1..46
FT /label= signal_peptide
FT /note= "SEC-dependent signal secretion sequence"
FT Protein 47..340
FT /label= Mature_protein

XX WO200204018-A2.
XX 17-JAN-2002.

XX 10-JUL-2001; 2001MO-US021717.
XX 10-JUL-2000; 2000US-0217646P.

XX (COLS) UNIV COLORADO STATE RES FOUND.
XX Orme IM, Belisle JT;
XX WPI; 2002-164602/21.
XX

PT Vaccine for boosting immunity to mycobacteria when administered in mid-
PT life in a subject who has been vaccinated in childhood with Bacillus
PT Calmette-Guerin, has purified proteins from mycobacterium tuberculosis.
XX

PS Claim 8; Page 17; 61pp; English.

XX The present sequence is that of the Mycobacterium tuberculosis (Mtb)
CC strain H37Rv gene Rv0129c product, designated FbpC2 Ag85C. This is one of
CC 31 immunodominant secreted or cytosolic Mtb proteins of strain H37Rv (see
CC AAM50729-59) discovered through the use of 2-dimensional liquid phase
CC electrophoresis coupled with an in vitro interferon-gamma assay and
CC liquid chromatography-mass spectrometry. The immunogens stimulate a
CC strong interferon-gamma response from T cells of M. tuberculosis infected
CC mice. The invention provides vaccine compositions for boosting immunity
CC to mycobacteria when administered in mid-life to a subject who has been
CC vaccinated neonatally or in early childhood with BCG and in whom
CC protective immunity has waned. The vaccine compositions comprise 1 or
CC more of the 31 purified immunogenic proteins, when used as immunogens,
CC the secreted Mtb proteins lack the secreted signal sequence. A preferred
CC protein is Ag85A (see AAM50759), the secreted product of the Rv3084v gene

XX Sequence 340 AA;

Query Match 34.8%; Score 564.5; DB 5; Length 340;
Best Local Similarity 41.6%; Pred. No. 1.3e-36;
Matches 131; Conservative 38; Mismatches 119; Indels 27; Gaps 8;

QY 4 RSA-----LRALMIA-----LSFGL-----GGVAVAEPTAKAPYENIMVPSPSMGNDIPV 52
DB 11 RSNATTLPRRLAIAMGAVLVYGLVGTGPGPATNAGAFSRGPLEVEYIQVPSASMGNDIKY 70

QY 53 AFLAGPRAVYLLDAFNAGPDVSNWVTAGNANTLAGKGISVVAAPAGAYSMNTNEQDG 112
DB 71 QFGGGGPHAVYLLDGLRAQDDYNGMDINTPAFEYYSQGLSVIMPVGQSSFYTDWYOPS 130

QY 113 SK-----QMDTFLSAELPDMILAANRGLAPGHAIVGAAGGYGAMALAAFPDRRGA 165
DB 131 QSNQNTYTKMETFLTRMPAMLOANKGVSPTGNAVGLSMGSGSALITLAATYPOQFPYA 190

QY 166 GSWGFLYPSNTTNGAIAAGMOQFGVDNCGMKGAPQLGRWKMHPWYHASTLLAQNTR 225
DB 191 ASLSGFLNPSBGWMPFLIGLANNDSGYNANSMWGPSDPAMKRNPMVQIPRLVANTR 250

QY 226 VWWV-----SPTNPGASD-PAAMI-GQTAEAMGNSRMFYNOYRSVGHNGHFDPPASGNDG 279
DB 251 IWYCGNGTFPSDLGGDNIPAKFLLEGTL---LRTNQTFRDYAADGGNGVFNFPNGTHS 307

QY 280 WGSNAPQLGAMSGDI 294
DB 308 WPYMNEQLVAMKADI 322

RESULT 14
AAR26167
ID AAR26167 standard; protein; 340 AA.

XX AAR26167;
XX 25-MAR-2003 (revised)
XX 02-FEB-1993 (first entry)

XX 85-C.
XX Immunise; antibody; vaccine; PCR; antisera; amplify.
XX Mycobacterium tuberculosis.

XX Key
XX Location/Qualifiers

FT Peptide 1..46
FT /label= signal_peptide
FT Protein 47..340
FT /label= Mature_peptide
XX

PN EP499003-A1.
 XX 19-AUG-1992.
 XX 14-FEB-1991; 91EP-00400388.
 XX 14-FEB-1991; 91EP-00400388.
 XX (INNO-) INNOGENETICS NV SA.
 XX Content J, De Wit L, De Bryn J;
 XX WPI, 1992-277793/34.
 DR N-PSDB; AAQ27378.
 XX
 PT Recombinant peptide(s) and their nucleic acids - for diagnosing
 PT tuberculosis and as a vaccine against tuberculosis.
 XX
 PS Disclosure; Fig 1; 48pp; English.
 XX
 CC The sequence given is 85-C from *Mycobacterium tuberculosis*. The 85-C gene
 CC was amplified from genomic DNA using the primer sequences given in
 CC AAQ27373-4. The 85-C gene was sequenced and the amino acid sequence
 CC determined. Various peptide motifs were identified and could then be used
 CC in the production of vaccines for immunisation against tuberculosis. The
 CC peptides may be used to raise antisera or antibodies against
 CC tuberculosis. (Updated on 25-MAR-2003 to correct PN field.)
 XX
 SQ Sequence 340 AA:
 Query Match 34.6%; Score 561.5; DB 2; Length 340;
 Best Local Similarity 41.3%; Pred. No. 2.3e-36;
 Matches 130; Conservative 39; Mismatches 119; Indels 27; Gaps 8;
 QY 4 RSR---LRLALMTAA---LSFGU---GGVAVAABPTAAAPYENLMVPSMGRDIPV 52
 DB 11 RSAATTLPRVVALAAMGAVLVYGLVGTGGPATGAASRGLPVRYLQVPSASGRDIXV 70
 QY 53 AFLAGPHAVYLLDAFNAGPVSVMWTAGNMMTLAKGISVVAAGAAVSMYTNWEDG 112
 DB 71 QFGGGPHAVYLLDGLRAQDDYNGMDINTPAFEYVYSGLSVIMPVGQSGSFYTDWYQPS 130
 QY 113 SK-----QMDTFLSALPDWTLAANRGLAPGHAAGAAGGAGMALAAFHDRGFPA 165
 DB 131 QSNQONTYTKMETFLTEMPAWLQANKGVSPYGAAGVGLSMGSGSALLILAAYYPOQFPYA 190
 QY 166 GSNMGLYPSNTNTNGAIAAGMOQFGVDNNGMGAQLGRWKHNDPMVHSLAQNTR 225
 DB 191 ASLSGFLNPSBGMWPTLIGLAMNDGGYNANSMGPPSSDPAMKNDPMVOIPRLVANNTR 250
 QY 226 VMTW----SPTNPGASD-PAAMI-GOTAEAMGNSRMFYNOYRSVYGNHGFDPASGNG 279
 DB 251 IMVYCGNGTSPSDGSDNIPAKFLBGLT---LRTNQTRDYAADGNGVFNFPNGTHS 307
 QY 280 WGSNAPQLGAMSGDI 294
 DB 308 WPTWNEQLVAMKADI 322
 RESULT 15
 AAAY14857
 ID AAAY14857 standard; protein; 340 AA.
 XX
 XX AAAY14857;
 DB 25-OCT-1999 (first entry)
 XX
 XX Antigen 85C protein from *M. tuberculosis*.
 KW *Mycobacterium vaccae* protein; antigen; T cell activation; cytokine;
 KW dendritic cell maturation; infectious disease; immune disorder; cancer;
 KW respiratory system; mycobacterial infection; allergy; tuberculosis;
 KW leprosy; sarcoidosis; lung cancer; asthma; skin disorder; psoriasis;

KW dermatitis; eczema; alopecia areata; skin cancer; basal carcinoma;
 KW squamous cell carcinoma; melanoma.
 XX
 OS *Mycobacterium tuberculosis*.
 XX
 XX W09932634-A2.
 XX
 XX 01-JUL-1999.
 XX
 XX 23-DEC-1998; 98MO-NZ000189.
 XX
 XX 23-DEC-1997; 97US-00996624.
 XX 23-DEC-1997; 97US-00997080.
 XX 23-DEC-1997; 97US-00997362.
 XX 11-JUN-1998; 98US-00095855.
 XX 17-SEP-1998; 98US-00156181.
 XX 04-DEC-1998; 98US-00205426.
 XX
 XX (GENE-) GENESIS RES & DEV CORP LTD.
 XX
 XX Tan P, Watson J, Visser BS, Skinner MA, Prestidge RJ;
 XX WPI; 1999-430163/36.
 XX
 XX Enhancing immune response to an antigen.
 XX
 PS Example 11; Page 160; 243pp; English.
 XX
 CC The invention provides heat-killed *Mycobacterium vaccae*, or recombinant
 CC *M. vaccae* proteins. The *M. vaccae* proteins may be employed to activate T
 CC cells and natural killer cells, to stimulate the production of cytokines,
 CC to enhance the expression of co-stimulatory molecules on dendritic cells
 CC and monocytes, and to enhance dendritic cell maturation and function. The
 CC proteins can be expressed by standard recombinant methodology.
 CC
 CC Pharmaceutical compositions comprising the proteins or nucleic acid
 CC sequences encoding the proteins can be used for the treatment,
 CC prevention, and detection of disorders including infectious diseases,
 CC immune disorders and cancer. In particular, the compounds and methods are
 CC used for treatment of diseases of the respiratory system, such as
 CC mycobacterial infections, asthma, allergies, tuberculosis, leprosy,
 CC sarcoidosis and lung cancers, and disorders of the skin such as
 CC psoriasis, atopic dermatitis, eczema, allergic contact dermatitis,
 CC alopecia areata, and skin cancers such as basal carcinoma, squamous cell
 CC carcinoma and melanoma
 XX
 SQ Sequence 340 AA:
 Query Match 34.6%; Score 561.5; DB 2; Length 340;
 Best Local Similarity 41.3%; Pred. No. 2.3e-36;
 Matches 130; Conservative 39; Mismatches 119; Indels 27; Gaps 8;
 QY 4 RSR---LRLALMTAA---LSFGU---GGVAVAABPTAAAPYENLMVPSMGRDIPV 52
 DB 11 RSAATTLPRVVALAAMGAVLVYGLVGTGGPATGAASRGLPVRYLQVPSASGRDIXV 70
 QY 53 AFLAGPHAVYLLDAFNAGPVSVMWTAGNMMTLAKGISVVAAGAAVSMYTNWEDG 112
 DB 71 QFGGGPHAVYLLDGLRAQDDYNGMDINTPAFEYVYSGLSVIMPVGQSGSFYTDWYQPS 130
 QY 113 SK-----QMDTFLSALPDWTLAANRGLAPGHAAGAAGGAGMALAAFHDRGFPA 165
 DB 131 QSNQONTYTKMETFLTEMPAWLQANKGVSPYGAAGVGLSMGSGSALLILAAYYPOQFPYA 190
 QY 166 GSNMGLYPSNTNTNGAIAAGMOQFGVDNNGMGAQLGRWKHNDPMVHSLAQNTR 225
 DB 191 ASLSGFLNPSBGMWPTLIGLAMNDGGYNANSMGPPSSDPAMKNDPMVOIPRLVANNTR 250
 QY 226 VMTW----SPTNPGASD-PAAMI-GOTAEAMGNSRMFYNOYRSVYGNHGFDPASGNG 279
 DB 251 IMVYCGNGTSPSDGSDNIPAKFLBGLT---LRTNQTRDYAADGNGVFNFPNGTHS 307
 QY 280 WGSNAPQLGAMSGDI 294

Db 308 WPTYNEQLVAMKADI 322

Search completed: April 14, 2006, 17:18:37
Job time : 117.851 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocelebration Ltd.

OM protein - protein search, using SW model

Run on: April 14, 2006, 17:19:08 ; Search time 17.5377 Seconds
(without alignments)
1640.866 Million cell updates/sec

Title: US-10-620-246-42

Perfect score: 1621

Sequence: 1 MKGRSALRLALMTALSLFGD.....WGSWAPQLGAMSGDIVGAIR 239

Scoring table: BL0SUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1616	99.7	299	2 G70887	fbpC1 protein prec
2	1251	77.2	301	2 B69821	antigen 85C, mycol
3	923.5	57.0	220	2 S32111	MPT51 protein - My
4	564.5	34.8	340	2 D70615	antigen fbpC2 - My
5	548	33.8	333	2 S32114	85C protein - Myco
6	532	32.8	338	2 S10326	alpha-antigen A, e
7	532	32.8	338	2 H70887	32K antigen fbpA p
8	521	32.1	330	2 A86921	antigen 85A, mycol
9	519	32.0	325	2 C70516	probable fbpB prot
10	513	31.6	325	2 S29663	antigen 85-B precu
11	512	31.6	323	2 A32348	alpha-antigen B pr
12	504.5	31.1	325	2 A37185	alpha-antigen prec
13	490.5	30.3	330	2 JN0897	alpha-antigen prec
14	490.5	30.3	330	2 S32773	alpha-antigen prec
15	478.5	29.5	327	2 G87162	antigen 85A, mycol
16	477.5	29.5	327	2 S20038	antigen 85-B - Myc
17	451.5	27.9	327	2 S20038	fibronectin-bindin
18	328.5	20.3	188	2 S32107	85A protein - Myco
19	260	16.0	657	2 S25184	cepi protein - Cor
20	206.5	12.7	342	2 T29115	hypothetical prote
21	174	10.7	456	2 D70772	hypothetical prote
22	125	7.5	489	2 B70619	hypothetical prote
23	122	7.5	322	2 A82822	esterase/lipase [i
24	122	7.5	322	2 E97600	tributylin esteras
25	121.5	7.5	783	2 G70824	hypothetical glyci
26	117	7.2	370	2 G70822	hypothetical glyci
27	113.5	7.0	414	2 S39530	poly(3-hydroxybuty
28	113.5	7.0	414	2 B70812	hypothetical glyci
29	111	6.8	1070	2 A10484	probable autotrans

30	110	6.8	461	2 F70571	hypothetical glyci
31	110	6.8	890	2 T35237	probable secreted
32	110	6.8	2639	2 T31328	fibron - Chinese
33	109.5	6.8	290	2 T23416	hypothetical prote
34	108	6.7	366	2 F70618	probable pncA pro
35	107.5	6.6	594	2 G70545	hypothetical glyci
36	107.5	6.6	606	2 H70816	hypothetical glyci
37	107.5	6.6	806	2 B49938	hupu protein - Rho
38	107.5	6.6	1053	2 B70987	probable pps prote
39	106.5	6.6	259	2 R95071	tributylin esteras
40	106.5	6.6	259	2 C97939	tributylin esteras
41	106.5	6.6	684	2 T36771	probable integral
42	106.5	6.6	1643	2 D71630	outer membrane pro
43	106	6.5	496	2 H70839	hypothetical glyci
44	106	6.5	543	2 F70726	hypothetical glyci
45	105.5	6.5	312	2 A70708	hypothetical prote

ALIGNMENTS

RESULT 1

G70887
fbpC1 protein precursor - Mycobacterium tuberculosis
N/Alternate names: major antigen MPT51
C/Species: Mycobacterium tuberculosis
C/Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C/Accession: G70887; A60274

R/Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, J.; Connor, R.; Davies, R.; Devlin, K.; Feltham, T.; Gentles, S.; Hamlin, N.; Holroyd, S. Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S. Nature 393, 537-544, 1998

A/Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G. A/Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome A/Reference number: A70500; MUID:98255987; PMID:9634230

A/Accession: G70887
A/Status: nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA

A/Residues: 1-299 <COL>

A/Cross-references: UNIPROT:Q48923; UNIPARC:UPI0000125437; GB:AL022076; GB:AL123456; NI

A/Experimental source: Strain H37Rv

R/Nagal, S.; Wiker, H.G.; Harboe, M.; Kilmoto, M. Infect. Immun. 59, 372-382, 1991

A/Title: Isolation and partial characterization of major protein antigens in the cultur

A/Reference number: A60274; MUID:91099989; PMID:1898899

A/Accession: A60274

A/Status: preliminary

A/Molecule type: protein

A/Residues: 34-40, 'Y', '42-49', 'K', '51-58', '59-76', '77-78 <NAG>

A/Cross-references: UNIPARC:UPI0000175922; UNIPARC:UPI00001794FC; UNIPARC:UPI00001794FD

C/Genetics:

A/Gene: fbpC1

C/Superfamily: Mycobacterium avium alpha-antigen

F/34-299/Product: fbpC1 protein #status experimental <MAT>

Query Match 99.7%; Score 1616; DB 2; Length 299;

Best Local Similarity 99.7%; Pred. No. 2.4e-111;

Matches 298; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY	1	MKGRSALRLALMTALSLFGIGVAVAAEPTAKAPENLWVPSPSMGRIIPVAFLAGPH 60		1	MKGRSALRLALMTALSLFGIGVAVAAEPTAKAPENLWVPSPSMGRIIPVAFLAGPH 60
DB	1	MKGRSALRLALMTALSLFGIGVAVAAEPTAKAPENLWVPSPSMGRIIPVAFLAGPH 60		1	MKGRSALRLALMTALSLFGIGVAVAAEPTAKAPENLWVPSPSMGRIIPVAFLAGPH 60
QY	61	AVYLLDAFNAGPVSWMVTAKGNTLAKGISVAVAPAGAYSMYTTWEDDSKQMTFL 120		61	AVYLLDAFNAGPVSWMVTAKGNTLAKGISVAVAPAGAYSMYTTWEDDSKQMTFL 120
DB	61	AVYLLDAFNAGPVSWMVTAKGNTLAKGISVAVAPAGAYSMYTTWEDDSKQMTFL 120		61	AVYLLDAFNAGPVSWMVTAKGNTLAKGISVAVAPAGAYSMYTTWEDDSKQMTFL 120
QY	121	SATLPDWLANRGLAAGAAVGAAGCGYAMLAAPHPRFPAGSMGFLVPSNTTN 180		121	SATLPDWLANRGLAAGAAVGAAGCGYAMLAAPHPRFPAGSMGFLVPSNTTN 180
DB	121	SATLPDWLANRGLAAGAAVGAAGCGYAMLAAPHPRFPAGSMGFLVPSNTTN 180		121	SATLPDWLANRGLAAGAAVGAAGCGYAMLAAPHPRFPAGSMGFLVPSNTTN 180
QY	181	GALAAQMGQGVGVDTNGMGAPQLGSKWHDPPVHA SLAQNNTRWVMSPTNPGASDPA 240		181	GALAAQMGQGVGVDTNGMGAPQLGSKWHDPPVHA SLAQNNTRWVMSPTNPGASDPA 240
DB	181	GALAAQMGQGVGVDTNGMGAPQLGSKWHDPPVHA SLAQNNTRWVMSPTNPGASDPA 240		181	GALAAQMGQGVGVDTNGMGAPQLGSKWHDPPVHA SLAQNNTRWVMSPTNPGASDPA 240

Db 181 GATAAGGQFGVDYTNMGMAPOLGRKMKHDPVYHASLLAQNNTVRVWVSPTEGASDPA 240
 QY 241 AMIGQTAEAMGNSRMPFYNYRSGVGHNGHFPDPASGDNMGWMAPOLGAMSGDIVGAIR 299
 Db 241 AMIGQAAEMGNSRMPFYNYRSGVGHNGHFPDPASGDNMGWMAPOLGAMSGDIVGAIR 299

RESULT 2
 B66921
 antigen 85C, mycolyltransferase [imported] - Mycobacterium leprae
 C/Species: Mycobacterium leprae
 C/Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
 C/Accession: B66921
 R./Cole, S.T.; Eiglmeyer, K.; Parkhill, J.; James, K.D.; Thomson, N.R.; Wheeler, P.R.; Ho
 R.; Davies, R.M.; Devlin, K.; Duthoy, S.; Feltwell, T.; Fraser, A.; Hamlin, N.; Holroyd,
 eam, M.A.; Rutherford, K.M.
 Nature 409, 1007-1011, 2001
 A/Authors: Rutter, S.; Seeger, K.; Simon, S.; Simmonds, M.; Skelton, J.; Squares, R.; Sc
 A/Title: Massive gene decay in the leprosy bacillus.
 A/Reference number: A86909; PMID:21128732; PMID:11234002
 A/Accession: B66921
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-301 <STO>
 A/Cross-references: UNIPROT:Q05868; UNIPARC:UPI000012P4B7; GB:AL450380; NID:q13092480; F
 C/Genetics:
 A/Gene: fdpC
 C/Superfamily: Mycobacterium avium alpha-antigen

Query Match 77.2%; Score 1251; DB 2; Length 301;
 Best Local Similarity 77.5%; Pred. No. 1.4e-84;
 Matches 234; Conservative 23; Mismatches 41; Indels 4; Gaps 2;

QY 1 MKGRSALLRALMTALSFGLGVAVAAPPT---AKAPYENIMVPSMGRDIPVAFLAG 57
 Db 1 MRGLSANNVRLCYBALAVGPAALAVLAGTAGNAKAGTESLWVPSNMGKRDIPVAFMAG 60
 QY 58 GPHAVYLLDAFNAGPDVSNVVTAGNANMTLAGKISVVAAPAGAYSMYTWEDGSKQMD 117
 Db 61 GPHAVYLLDAFNALDVSNNVTAGNMTTLAGRGISVVAAPAGAYSMYTWEDGSKQMD 120
 QY 118 TPLSASELPDLAANRGLAPGSHAAGVGAAGGYGAMLLAHPHPRFGASMSGFLYPSNT 177
 Db 121 TPLSSELPDLATKRGGLAPGSHAAGVGAAGGYGAMLLAHPHPRFGASLSGFLYPSST 180
 QY 178 TTNGATLAAGQFGVDYTNMGMAPOLGRKMKHDPVYHASLLAQNNTVRVWVSPTEGAS 237
 Db 181 NTNGATLAAGQFGVDYTNMGMAPOLGRKMKHDPVYHASLLAQNNTVRVWVSPTEGAS 239
 QY 238 DPAAMTIGQTAEAMGNSRMPFYNYRSGVGHNGHFPDPASGDNMGWMAPOLGAMSGDIVGA 297
 Db 240 DPAAMTIGQAAEMGNSRMPFYNYRSGVGHNGHFPDPASGDNMGWMAPOLGAMSGDIVGA 299

QY 298 IR 299
 Db 300 IR 301

RESULT 3
 S32111
 MPT51 protein - Mycobacterium leprae (fragment)
 C/Species: Mycobacterium leprae
 C/Date: 06-Feb-1995 #sequence_revision 06-Feb-1995 #text_change 09-Jul-2004
 C/Accession: S32111
 R./Rinke de Wit, T.F.; Bekelke, S.; Oeland, A.; Wielea, B.; Jansen, A.A.M.; Thole, J.B.R.
 submitted to the EMBL Data Library, March 1993
 A/Description: The M.leprae antigen 85 complex gene family: identification of the genes
 A/Reference number: S32107
 A/Accession: S32111
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-220 <RIN>
 A/Cross-references: UNIPROT:Q05868; UNIPARC:UPI000016FB12; EMBL:Z21949; NID:q287923; PDI

C/Genetics:
 A/Start codon: GTG
 C/Superfamily: Mycobacterium avium alpha-antigen

Query Match 57.0%; Score 923.5; DB 2; Length 220;
 Best Local Similarity 78.2%; Pred. No. 1e-60;
 Matches 172; Conservative 19; Mismatches 26; Indels 3; Gaps 1;

QY 1 MKGRSALLRALMTALSFGLGVAVAAPPT---AKAPYENIMVPSMGRDIPVAFLAG 57
 Db 1 MRGLSANNVRLCYBALAVGPAALAVLAGTAGNAKAGTESLWVPSNMGKRDIPVAFMAG 60
 QY 58 GPHAVYLLDAFNAGPDVSNVVTAGNANMTLAGKISVVAAPAGAYSMYTWEDGSKQMD 117
 Db 61 GPHAVYLLDAFNALDVSNNVTAGNMTTLAGRGISVVAAPAGAYSMYTWEDGSKQMD 120
 QY 118 TPLSASELPDLAANRGLAPGSHAAGVGAAGGYGAMLLAHPHPRFGASMSGFLYPSNT 177
 Db 121 TPLSSELPDLATKRGGLAPGSHAAGVGAAGGYGAMLLAHPHPRFGASLSGFLYPSST 180
 QY 178 TTNGATLAAGQFGVDYTNMGMAPOLGRKMKHDPVYHAS 217
 Db 181 NTNGATLAAGQFGVDYTNMGMAPOLGRKMKHDPVYHAS 220

RESULT 4
 D70615
 antigen fdpC2 - Mycobacterium tuberculosis (strain H37RV)
 N/Alternate names: antigen 85-C
 C/Species: Mycobacterium tuberculosis
 C/Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
 C/Accession: D70615; A43603; S15504
 R./Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon,
 R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.
 Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
 Nature 393, 537-544, 1998
 A/Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
 A/Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
 A/Reference number: A70500; PMID:98295987; PMID:9634230
 A/Accession: D70615
 A/Status: nucleic acid sequence not shown; translation not shown
 A/Molecule type: DNA
 A/Residues: 1-340 <COL>
 A/Cross-references: UNIPROT:P31953; UNIPARC:UPI0000042501; GB:Z92770; GB:AL123456; NID:
 A/Experimental source: strain H37RV
 A/Content: J.; de la Cueva-Lleria, A.; De Wit, L.; Vincent-Ley-Frebaule, V.; Ooms, J.; D
 Infect. Immun. 59, 3205-3212, 1991
 A/Title: The genes coding for the antigen 85 complexes of Mycobacterium tuberculosis an
 n of the gene coding for antigen 85-C of Mycobacterium tuberculosis.
 A/Reference number: A43603; MOID:91348869; PMID:1715324
 A/Accession: A43603
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-20, 'V', 22-340 <CON>
 A/Cross-references: UNIPARC:UPI000002DC31; GB:X57229; NID:q48827; PIDN:CAA40506.1; PID:
 C/Genetics:
 A/Gene: fdpC2
 C/Superfamily: Mycobacterium avium alpha-antigen

Query Match 34.8%; Score 564.5; DB 2; Length 340;
 Best Local Similarity 41.6%; Pred. No. 3.3e-34;
 Matches 131; Conservative 38; Mismatches 119; Indels 27; Gaps 8;

QY 4 RSA---LRLALMTA---LSFGL-----GGVAVAAPPTAKAPYENIMVPSMGRDIPV 52
 Db 11 RSAAATTLPRRLAIAAGAVLVYGLVGTGFPATAGFSRGLPVEYLQVPSASMGKDIK 70
 QY 53 AFLAGGPHAVYLLDAFNAGPDVSNVVTAGNANMTLAGKISVVAAPAGAYSMYTWEDG 112
 Db 71 QFGGQPHAVYLLDGLRAQDDVYNGWDINPAPFEYVYQSGLSVIMPVGGQGSFTDWYQPS 130
 QY 113 SK-----QWDITLASELPDLAANRGLAPGSHAAGVGAAGGYGAMLLAHPHPRFG 165

Db 131 QSGNQYTYKWEFTLTETPMALQANKGVSPTGNAAVGLSMGSSALITLAATYQOPRYA 190
 QY 166 GSNMGLPYPSNTTNGAIAAGMOQFGVDITNGMGAPOLGRWKMDPWHASLLAQNNTT 225
 Db 191 ASISGFLNPSBGMWPTLIGLAMDSSGYNNNSMWGSPSDPAWKENDPMVQIPRLVANNTR 250
 QY 226 WWTW---SPTNPGASD-PAAMI-GQTAEAMGNSRMFTYNOYRVSNGHGFDPASGDNG 279
 Db 251 IHWYCGNGTSPSDLGDDNI PAKFLEGLT---LRTNQTRDYTAADGGRNGVFNPPNGTHS 307
 QY 280 WGSMAPOLGAMSGDI 294
 Db 308 WPMYNEQLVAMKADI 322

RESULT 5

85C protein - Mycobacterium leprae
 S32114
 C/Species: Mycobacterium leprae
 C/Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Jul-2004
 C/Accession: S32114
 R/Inke de Wit, T.F.; Bekke, S.; Oeland, A.; Wietes, B.; Janson, A.A.M.; Thole, J.B.R.
 submitted to the EMBL Data Library, March 1993
 A/Description: The M. leprae antigen 85 complex gene family: identification of the genes
 A/Reference number: S32107
 A/Accession: S32114
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1333 <RIN>
 A/Cross-references: UNIPROT:Q05862; UNIPARC:UPI0000125061; EMBL:Z21951; NID:G287921; PIR
 C/Superfamily: Mycobacterium avium alpha-antigen

Query Match 33.8%; Score 548; DB 2; Length 333;
 Best Local Similarity 39.2%; Pred. No. 5.3e-33;

Matches 118; Conservative 41; Mismatches 128; Indels 14; Gaps 5;

QY 6 ALLRALMIALSFGGLGVAVAEPTAKAAYENIMVPSMGNDIPVAFLAGPH--AVYLLDA 65
 Db 24 AVIGTALMAGLVGVGDTALVAFSKFGLPVEYLQVPSMGNDIKIQFGGGGHAYLL 83
 QY 66 DAPNAPDVSNWVTAGNANTLAGKISVAPAGGAYSMITTNEQ--DSKQ----WDT 118
 Db 84 DGLRAQEDVNGWMDINTPAFEYTHSGLSVIMPGGQSSFSNMYQPSQNGQHYTKMET 143
 QY 119 FLSAEIPDMTLAARGLAPGHAAGVGAAGGVMALAFHPDRFGPAGSMGFLYPSNTT 178
 Db 144 FLTQENPSMLQANKNVLPTRAAVGLSMGSSALITLASYTPQOPRYAASISGFLNPSBGM 203
 QY 179 TNGAIAAGMOQFGVDITNGMGAPOLGRWKMDPWHASLLAQNNTTWTW---SPTNP 234
 Db 204 WPTMIGLAMDSSGYNNNSMWGSPSDPAWKENDPMVQIPRLVANNTRITWYCGNGANFEL 263
 QY 235 GASD-PAAMIGQTAEAMGNSRMFTYNOYRVSNGHGFDPASGDNGSMAPOLGAMSGD 293
 Db 264 GGNINIPAKFL--ESLTLSTNEIFONTYAAGSGRNGVFNPPNGTHSWPMNQOLVAMKPD 321
 QY 294 I 294
 Db 322 I 322

RESULT 6

S10326
 alpha-antigen A, extracellular - Mycobacterium bovis
 C/Species: Mycobacterium bovis
 C/Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 09-Jul-2004
 C/Accession: S10326; B41499
 R/de Wit, L.; de la Cuvelier, A.; Ooms, J.; Content, J.
 Nucleic Acids Res. 18, 3995, 1990
 A/Title: Nucleotide sequence of the 32 kDa-protein gene (antigen 85 A) of Mycobacterium
 A/Reference number: S10326; MUID:90326531; PMID:2197602
 A/Accession: S10326
 A/Status: preliminary

A/Molecule type: DNA
 A/Residues: 1-338 <WIT>
 A/Cross-references: UNIPROT:P17944; UNIPARC:UPI000002CE66; EMBL:X53034; NID:G44165; PID
 R/Wiker, H.G.; Sletten, K.; Nagai, S.; Harboe, M.
 Infect. Immun. 58, 272-274, 1990
 A/Title: Evidence for three separate genes encoding the proteins of the mycobacterial a
 A/Reference number: A41499; MUID:90093478; PMID:2403534
 A/Accession: B41499
 A/Status: preliminary
 A/Molecule type: protein
 A/Residues: 44-82 <WIK>
 A/Cross-references: UNIPARC:UPI000002PA00
 C/Superfamily: Mycobacterium avium alpha-antigen

Query Match 32.8%; Score 532; DB 2; Length 338;
 Best Local Similarity 40.1%; Pred. No. 8e-32;

Matches 121; Conservative 37; Mismatches 126; Indels 18; Gaps 5;

QY 14 AALSFGI---GGVAVAEPTAKAAYENIMVPSMGNDIPVAFLAGPH--AVYLLDA 67
 Db 25 AALVSGLVGAVGATATGAFSRPGLPVEYLQVPSMGNDIKIQFGGGANSPALYLLDG 84
 QY 68 FNAQPDVSNWVTAGNANTLAGKISVAPAGGAYSMITTNEQD-----SKQNTFL 120
 Db 85 LRAQDPFSGMDINTPAFEYTHSGLSVIMPGGQSSFSNMYQPSQNGQHYTKMETFL 144
 QY 121 SAEIPDMTLAARGLAPGHAAGVGAAGGVMALAFHPDRFGPAGSMGFLYPSNTTN 180
 Db 145 TSLIPGFLQANKNVLPTRAAVGLSMGSSALITLHYPOQFYAGMNSGLDPSQMG 204
 QY 181 GATAAGMOQFGVDITNGMGAPOLGRWKMDPWHASLLAQNNTTWTW---SPTNPGA 236
 Db 205 TLIGLAMDAGGRKASPMGPKEDPAMQGRNDPLNNGKLIANNTRVWYVYCGNGKPSDLGG 264
 QY 237 SDPAAMIGQTAEAMGNSRMFTYNOYRVSNGHGFDPASGDNGSMAPOLGAMSGIVG 296
 Db 265 NNLPKFLSEGFVTSNIK-FQDAYNAGGNGVFPDPDSTHMEYGAQILAMKPDLO 323
 QY 297 AI 298
 Db 324 AL 325

RESULT 7

H70887
 32K antigen fmpa precursor - Mycobacterium tuberculosis (strain H37RV)
 N/Alternate names: 30K native antigen; major protein antigen MPT45
 C/Species: Mycobacterium tuberculosis
 C/Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
 C/Accession: H70887; A37024; G60274; A54318
 R/Cole, S.T.; Broesch, R.; Parhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon,
 J.; Connor, R.; Davies, R.; Devlin, K.; Fellwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.
 Rajadream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
 Nature 393, 537-544, 1998
 A/Authors: Squares, R.; Suleston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
 A/Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
 A/Reference number: A70500; MUID:9825987; PMID:9634230
 A/Accession: H70887
 A/Status: nucleic acid sequence not shown; translation not shown
 A/Molecule type: DNA
 A/Residues: 1-338 <COL>
 A/Cross-references: UNIPROT:P17944; UNIPARC:UPI000002CE66; GB:AL022076; GB:AL123456; NI
 A/Experimental sources: strain H37RV
 R/Bortmann, M.; De Wit, L.; Volckaert, G.; Ooms, J.; De Bruyn, J.; Huygen, K.; Van Voo
 Infect. Immun. 57, 3123-3130, 1989
 A/Title: Cloning, sequence determination, and expression of a 32-kilodalton-protein gen
 A/Reference number: A37024; MUID:89379378; PMID:2506131
 A/Accession: A37024
 A/Molecule type: DNA
 A/Residues: 1-23, 25, 'R', 27-130, 'R', 132-291, 'R', 293-323, 'HWVRRPT', 332-333, 335-338 <BOR>
 A/Cross-references: UNIPARC:UPI00001794PB; GB:M27016; GB:X53898
 R/Nagai, S.; Wiker, H.G.; Harboe, M.; Kikimoto, M.
 Infect. Immun. 59, 372-382, 1991


```

Db      93  INTPAEWYQSGSLIYMPVGGSSFYSDKTSACGACGCTTYKMTFLITSELPQWLSAN 152
      132  RGLAPGHAHVGAAGQGYGAMALAAFHDPDRFGAGNSGFLYPSNTTNGAIAAGMOQFG 191
      153  RAYKPTGSAIIGLSMAGSSAMILAAHYHPOQIYAGISALIDPBGQMGPSLIGLAMDAG 212
Qy      192  GVDITNGMGAPQLGRWTHMDPWVHASLLAQNTRVWYV---SPTN-PGASDPAAMIGQT 246
      213  GYRAADWMPSSDPAMERNPTQOIPLVANNTRLMVCNGTGNELGANNIPAEFLBNF 272
Qy      247  AEMAGNSRMFYNYQRYSGHNGHDPFASGDNKGSAWAPOLGAMSGDIYVAI 298
      273  VRS--SNLKFDQDAYNAAGHNAVFNFPNGTHSWETYGAGQILNMMKGLDQSSL 322

RESULT 10
S29663
antigen 85-B precursor - Mycobacterium bovis
C/Species: Mycobacterium bovis
C/Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Jul-2004
C/Accession: S29663
R/de Wit, L.; Palou, M.; Content, J.
submitted to the EMBL Data Library, December 1991
A/Reference number: S29663
A/Accession: S29663
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-325 <MT>
A/Cross-references: UNIPROT:P12942; UNIPARC:UPI000016F90E; EMBL:X62397; NID:g44167; PIDD:
C/Superfamily: Mycobacterium avium alpha-antigen

Query Match      31.6%; Score 513; DB 2; Length 325;
Best Local Similarity 38.4%; Pred. No. 1.9e-30;
Matches 112; Conservative 39; Mismatches 125; Indels 16; Gaps 5;

Qy      21  GGVAAAEPTAKAPYENIMVPSPSMGRIIPVAFLAGPH--AVYLLDAFAGPDVSNV 78
      33  GGAATGAFSPGIPVEYILQVPSPSMGRIKVFQSGGNSPAYVLLDGLRAQDDYNGWD 92
Db      79  TAGNANNTLAGKISVVAAPAGAVSMYTNMEODG-----SKQMDTFLSAPDMLAAN 131
      93  INTPAEWYQSGSLIYMPVGGSSFYSDKTSACGACGCTTYKMTFLITSELPQWLSAN 152
Qy      132  RGLAPGHAHVGAAGQGYGAMALAAFHDPDRFGAGNSGFLYPSNTTNGAIAAGMOQFG 191
      153  RAYKPTGSAIIGLSMAGSSAMILAAHYHPOQIYAGISALIDPBGQMGPSLIGLAMDAG 212
Qy      192  GVDITNGMGAPQLGRWTHMDPWVHASLLAQNTRVWYV---SPTN-PGASDPAAMIGQT 246
      213  GYRAADWMPSSDPAMERNPTQOIPLVANNTRLMVCNGTGNELGANNIPAEFLBNF 272
Qy      247  AEMAGNSRMFYNYQRYSGHNGHDPFASGDNKGSAWAPOLGAMSGDIYVAI 298
      273  VRS--SNLKFDQDAYNAAGHNAVFNFPNGTHSWETYGAGQILNMMKGLDQSSL 322

RESULT 11
A32348
alpha-antigen B precursor, extracellular - Mycobacterium bovis
N/Alternate names: 32k antigen
C/Species: Mycobacterium bovis
C/Date: 21-May-1990 #sequence_revision 21-May-1990 #text_change 09-Jul-2004
C/Accession: A32348; A41499; A60278
R/Matano, K.; Yamaguchi, R.; Yamazaki, A.; Tasaka, H.; Yamada, T.
J. Bacteriol. 170, 3847-3854, 1988
A/Title: Cloning and expression of the Mycobacterium bovis BCG gene for extracellular alpha-
A/Reference number: A32348; MUID:88314872; PMID:2842287
A/Accession: A32348
A/Molecule type: DNA
A/Residues: 1-323 <MT>
A/Cross-references: UNIPROT:P12942; UNIPARC:UPI000012505B; GB:M21839; NID:g149935; PIDD:
R/Wiker, H.G.; Sletten, K.; Nagai, S.; Harboe, M.
Infect. Immun. 58, 272-274, 1990

```

```

A/Title: Evidence for three separate genes encoding the proteins of the mycobacterial a
A/Reference number: A41499; MUID:90093478; PMID:2403534
A/Accession: A41499
A/Molecule type: protein
A/Residues: 41-85 <WT>
A/Cross-references: UNIPARC:UPI00001794F7
R/Filiis, T.; Costopoulos, C.; Radford, A.J.; Bactic, A.; Wood, P.R.
Infect. Immun. 59, 800-807, 1991
A/Title: Purification and characterization of major antigens from a Mycobacterium bovis
A/Reference number: A60278; MUID:91147217; PMID:1900061
A/Accession: A60278
A/Molecule type: protein
A/Residues: 41-59 <RT>
A/Cross-references: UNIPARC:UPI00001794F8
C/Superfamily: Mycobacterium avium alpha-antigen
F:1-40/Domain: signal sequence #status predicted <SIG>
F:41-323/Product: alpha antigen, extracellular #status experimental <MAT>

Query Match      31.6%; Score 512; DB 2; Length 323;
Best Local Similarity 38.7%; Pred. No. 2.2e-30;
Matches 113; Conservative 39; Mismatches 122; Indels 18; Gaps 6;

Qy      21  GGVAAAEPTAKAPYENIMVPSPSMGRIIPVAFLAGPH--AVYLLDAFAGPDVSNV 78
      33  GGAATGAFSPGIPVEYILQVPSPSMGRIKVFQSGGNSPAYVLLDGLRAQDDYNGWD 92
Db      79  TAGNANNTLAGKISVVAAPAGAVSMYTNMEODG-----SKQMDTFLSAPDMLAAN 131
      93  INTPAEWYQSGSLIYMPVGGSSFYSDKTSACGACGCTTYKMTFLITSELPQWLSAN 152
Qy      132  RGLAPGHAHVGAAGQGYGAMALAAFHDPDRFGAGNSGFLYPSNTTNGAIAAGMOQFG 191
      153  RAYKPTGSAIIGLSMAGSSAMILAAHYHPOQIYAGISALIDPBGQMGPSLIGLAMDAG 210
Qy      192  GVDITNGMGAPQLGRWTHMDPWVHASLLAQNTRVWYV---SPTN-PGASDPAAMIGQT 246
      211  GYRAADWMPSSDPAMERNPTQOIPLVANNTRLMVCNGTGNELGANNIPAEFLBNF 270
Qy      247  AEMAGNSRMFYNYQRYSGHNGHDPFASGDNKGSAWAPOLGAMSGDIYVAI 298
      271  VRS--SNLKFDQDAYNAAGHNAVFNFPNGTHSWETYGAGQILNMMKGLDQSSL 320

RESULT 12
A37185
alpha-antigen precursor - Mycobacterium kanasii
C/Species: Mycobacterium kanasii
C/Date: 31-Jan-1992 #sequence_revision 31-Jan-1992 #text_change 09-Jul-2004
C/Accession: A37185
R/Matano, K.; Yamaguchi, R.; Yamazaki, A.; Tasaka, H.; Terasaka, K.; Yamada, T.
Infect. Immun. 58, 550-556, 1990
A/Title: Cloning and expression of the gene for the cross-reactive alpha antigen of Myc
A/Reference number: A37185; MUID:90129315; PMID:2404875
A/Accession: A37185
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-325 <MAT>
A/Cross-references: UNIPROT:P21160; UNIPARC:UPI000012505D; GB:M53897
C/Superfamily: Mycobacterium avium alpha-antigen

Query Match      31.1%; Score 504.5; DB 2; Length 325;
Best Local Similarity 36.7%; Pred. No. 7.9e-30;
Matches 115; Conservative 47; Mismatches 132; Indels 19; Gaps 7;

Qy      3  GSEALLPALMTALS--FGL-GGVAAAEPTAKAPYENIMVPSPSMGRIIPVAFLAGP 59
      12  GRRLVLGAATAAALPLGLVGLAGAAATAGAFSRGLPEYILQVPSAAMGRSITKQVFOGGD 71
Db      60  H--AVYLLDAFAGPDVSNVVTAGNANNTLAGKISVVAAPAGAVSMYTNMEODG----- 112
      72  NSPAVYLLDGLRAQDDYNGWDITPAPEWYQSGSLIYMPVGGSSFYSDKTSACGAK 131
Qy      113  --SKQMDTFLSAPDMLAANRGLAPGHAHVGAAGQGYGAMALAAFHDPDRFGAGMSG 170

```

Db 132 CTTYKKEITLTSBELPQWLSANRSVKTGSAVGI SNAGSSALITLSYVHPQOFIYAGSLSL 191
 QY 171 FLYPSNTTNGAIAAGMOQFGVDITNGMGAPQLGRMKHDPVWVHSLLAQNTRVWV 229
 Db 192 LMDPSQGMGSPSLIGLAMGDAAGYKADAMWGSPSSDPAMQNRDPSLHIBELVANNTRIMLYC 251
 QY 230 ---SPTN-PGASDPAAIMQGTAEAMGNSRMFYNQYRSVGGNHGHPDPFASGDNMGSMAP 285
 Db 252 GNGTPEBELGAVNPAFFLENFVRS--SNLKFQDAYNAAGHNNAVFNLDNNGTHSWEYWG 309
 QY 286 QLGAMSGDIVGAI 298
 Db 310 QLNAMKPDLOASL 322

RESULT 13

alpha-antigen precursor - Mycobacterium intracellulare
 N:Alternate names: Ag85B
 C:Species: Mycobacterium intracellulare
 C>Date: 19-May-1994 #sequence_revision 19-May-1994 #text_change 09-Jul-2004
 C/Accession: JN0897
 R:Kleitura, H.; Ohara, N.; Matsuo, T.; Tasaka, H.; Kobayashi, K.; Yamada, T.
 Biochem. Biophys. Res. Commun. 196, 1466-1473, 1993
 A>Title: Cloning, sequencing and expression of the gene for alpha antigen from Mycobacterium intracellulare
 A:Reference number: JN0897, M01D:94071912, PMID:8250904
 A:Accession: JN0897
 A:Molecule type: DNA
 A:Residues: 1-330 <KIT>
 A:Cross-references: UNIPROT:Q49575; UNIPARC:UP1000012505C
 C:Comment: This protein is one of antigen 85.
 C:Comment: This protein is an important stimulant of cellular and humoral immunity and has been shown to be a major component of the cell wall.
 C:Superfamily: Mycobacterium avium alpha-antigen
 F:1-40/Domain: signal sequence #status predicted <SIG>
 F:41-330/Product: alpha-antigen #status predicted <PAA>

Query Match 30.3%; Score 490.5; DB 2; Length 330;
 Best Local Similarity 37.6%; Pred. No. 8, 6e-29;
 Matches 118; Conservative 41; Mismatches 134; Indels 21; Gaps 7;
 QY 3 GRSALLRALMIALSFGI---GGVAVAAPETAKAPYENLWVPSPMGRDIPVAFIAG- 57
 Db 12 GRLIVGAAGAAATLP-GLIGIAGGAATANAFSRPGLPEVEYLQVPSAGMGHDIKVQFQSGG 70
 QY 58 -GHNATYLLDAFPAAGDVSWWVTRAGNANTLAKGISVVA-PAGAVSMYTNWED----- 111
 Db 71 NGSFAYVLLDGLRAQDDYNGMDINTPAFEWYQSGLSVMPVGQSSFYADWYQPAACGKA 130
 QY 112 --GSKOMDFLSAELPDMILANRGLAPGHAAGAAGYGAAMALAAFPDRFGFAGSMG 169
 Db 131 GCGTYKMDITFUTSELFOYLAISNGKYSTGSAVGISMSGSAHILAVNHENQVYVAGSL 190
 QY 170 GFLPSNTTNGAIAAGMOQFGVDITNGMGAPQLGRMKHDPVWVHSLLAQNTRVWV 229
 Db 191 ALLDPGQMGPSLIGLAMGDAAGYKADAMWGSPSSDPAMQNRDPSLHIBELVANNTRIMLYC 250
 QY 230 ---SPTN-PGASDPAAIMQGTAEAMGNSRMFYNQYRSVGGNHGHPDPFASGDNMGSMAP 285
 Db 251 CGNGTPEBELGAVNPAFFLENFVRS--SNLKFQDAYNAAGHNNAVFNLDNNGTHSWEYWG 309
 QY 285 POLGAMSGDIVGAI 298
 Db 309 AQLNAMKPDLOASL 322

RESULT 14

alpha-antigen - Mycobacterium avium
 C/Species: Mycobacterium avium
 C/Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 31-Dec-2004
 C/Accession: S32773
 R:Ohara, N.; Matsuo, K.; Yamaguchi, R.; Yamazaki, A.; Tasaka, H.; Yamada, T.

submitted to the EMBL Data Library, December 1991
 A>Description: Nucleotide sequence of the alpha-antigen gene of Mycobacterium avium.
 A:Reference number: S32773
 A:Accession: S32773
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-330 <OHA>
 A:Cross-references: UNIPROT:Q8L371; UNIPROT:Q8V082; UNIPARC:UP100001794F6; EMBL:X63437

Query Match 30.3%; Score 490.5; DB 2; Length 330;
 Best Local Similarity 37.4%; Pred. No. 8, 6e-29;
 Matches 117; Conservative 42; Mismatches 135; Indels 19; Gaps 7;

QY 3 GRSALLRALMIALSFGI---GGVAVAAPETAKAPYENLWVPSPMGRDIPVAFIAG-- 57
 Db 12 GRLIVGAAGAAATLPGLIGIAGGAATANAFSRPGLPEVEYLQVPSAGMGHDIKVQFQSGGN 71
 QY 58 GHNATYLLDAFPAAGDVSWWVTRAGNANTLAKGISVVA-PAGAVSMYTNWED----- 111
 Db 72 GSPAVVLLDGLRAQDDYNGMDINTPAFEWYQSGLSVMPVGQSSFYADWYQPAACGKA 131
 QY 112 -GSKOMDFLSAELPDMILANRGLAPGHAAGAAGYGAAMALAAFPDRFGFAGSMG 170
 Db 132 CTTYKKEITLTSBELPQWLSANRSVKTGSAVGI SNAGSSALITLSYVHPQOFIYAGSLSL 191
 QY 171 FLYPSNTTNGAIAAGMOQFGVDITNGMGAPQLGRMKHDPVWVHSLLAQNTRVWV- 229
 Db 192 LMDPSQGMGSPSLIGLAMGDAAGYKADAMWGSPSSDPAMQNRDPSLHIBELVANNTRIMLYC 251
 QY 230 ---SPTN-PGASDPAAIMQGTAEAMGNSRMFYNQYRSVGGNHGHPDPFASGDNMGSMAP 285
 Db 252 GNGTPEBELGAVNPAFFLENFVRS--SNLKFQDAYNAAGHNNAVFNLDNNGTHSWEYWG 309
 QY 286 QLGAMSGDIVGAI 298
 Db 310 QLNAMKPDLOASL 322

RESULT 15

antigen 85A, mycolyltransferase [imported] - Mycobacterium leprae
 C/Species: Mycobacterium leprae
 C/Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
 C/Accession: G87162
 R:Cole, S.T.; Eiglmeier, K.; Parish, J.; James, K.D.; Thomson, N.R.; Wheeler, P.R.; R. Davies, R.M.; Devlin, K.; Duthoy, S.; Felwell, T.; Fraser, A.; Hamlin, N.; Holroyd, eam, M.A.; Rutherford, K.M.
 Nature 409, 1007-1011, 2001
 A:Authors: Rutter, S.; Seeger, K.; Simon, S.; Simmonds, M.; Skelton, J.; Squares, R.; S.
 A>Title: Massive gene decay in the leprosy bacillus.
 A:Reference number: A86909; M01D:21128732; PMID:11234002
 A:Accession: G87162
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-327 <STO>
 A:Cross-references: UNIPROT:P31951; UNIPARC:UP1000012505E; GB:AL450380; M01D:913093650;
 C:Gene: fbpB
 C:Superfamily: Mycobacterium avium alpha-antigen

Query Match 29.5%; Score 478.5; DB 2; Length 327;
 Best Local Similarity 36.6%; Pred. No. 6, 4e-28;
 Matches 112; Conservative 40; Mismatches 133; Indels 21; Gaps 6;

QY 3 GRSALLRALMIALSFGI---GGVAVAAPETAKAPYENLWVPSPMGRDIPVAFIAG--GPH 60
 Db 18 GAATPLPSLISLA-----GGAATASAFSRPGLPEVEYLQVPSAGMGHDIKVQFQSGGN 72
 QY 61 AVYLLDAFPAAGDVSWWVTRAGNANTLAKGISVVA-PAGAVSMYTNWEDG-----S 113
 Db 73 AVYLLDGLRAQDDYNGMDINTSAPFEWYQSGLSVMPVGQSSFYADWYQPAACGKA 132
 QY 114 KOMDFLSAELPDMILANRGLAPGHAAGAAGYGAAMALAAFPDRFGFAGSMGFLY 173

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 14, 2006, 17:04:10 ; Search time 110.321 Seconds
(without alignments)
1912.171 Million cell updates/sec

Title: US-10-620-246-42

Perfect score: 1621
Sequence: 1 MGGRSALLRALMTALSLFGL.....WGSWAPQLGAMSGDIVGAIR 299

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Uniprot_05.80:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1616	99.7	299	1 MPT51_MYCBO	POA47 mycobacteri
2	1616	99.7	299	1 MPT51_MYCBO	POA46 mycobacteri
3	1328.5	82.0	300	2 Q74512_MYCPA	Q74512 mycobacteri
4	1251	77.2	301	1 MPT51_MYCLE	Q05868 mycobacteri
5	572.5	35.3	352	1 A85C_MYCAV	Q52972 mycobacteri
6	572.5	35.3	352	2 Q73037_MYCPA	Q73037 mycobacteri
7	569.5	35.1	352	2 Q8KX58_MYCPA	Q8KX58 mycobacteri
8	564.5	34.8	340	1 A85C_MYCBO	POA45 mycobacteri
9	564.5	34.8	340	1 A85C_MYCBO	POA44 mycobacteri
10	548	33.8	333	1 A85C_MYCLE	Q05862 mycobacteri
11	532	32.8	338	1 A85A_MYCBO	POA43 mycobacteri
12	532	32.8	338	1 A85A_MYCBO	POA42 mycobacteri
13	532	32.8	338	2 Q847N5_MYCBO	Q847N5 mycobacteri
14	532	32.8	338	2 Q6PU6_MYCBO	Q6PU6 mycobacteri
15	521	32.1	330	1 A85A_MYCLE	Q05861 mycobacteri
16	520	32.1	335	2 Q847N4_MYCBO	Q847N4 mycobacteri
17	519	32.0	325	1 A85B_MYCBO	PI2942 mycobacteri
18	519	32.0	325	1 A85B_MYCBO	PI2942 mycobacteri
19	504.5	31.1	325	1 A85B_MYCBO	PI2942 mycobacteri
20	492.5	30.3	330	2 Q732J4_MYCPA	Q732J4 mycobacteri
21	491.5	30.3	330	2 Q8VU82_MYCPA	Q8VU82 mycobacteri
22	490.5	30.3	330	1 A85B_MYCAV	Q06947 mycobacteri
23	490.5	30.3	330	1 A85B_MYCAV	Q06947 mycobacteri
24	489.5	30.2	330	2 Q8L371_MYCPA	Q8L371 mycobacteri
25	480.5	29.6	339	1 A85A_MYCBO	Q06052 mycobacteri
26	478.5	29.5	337	1 A85B_MYCLE	PI1951 mycobacteri
27	475	29.3	337	1 A85A_MYCUL	PI58248 mycobacteri
28	474	29.2	330	1 A85B_MYCSC	Q05037 mycobacteri
29	469.5	29.0	347	1 A85A_MYCAV	Q52956 mycobacteri
30	469	28.9	337	2 Q7X511_MYCNR	Q7X511 mycobacteri
31	461.5	28.5	347	2 Q8KX59_MYCPA	Q8KX59 mycobacteri

32	444.5	27.4	304	2 Q70E87_MYCPA	Q70E87 mycobacteri
33	358.5	22.1	345	2 Q523G6_NOCRA	Q523G6 nocardia fa
34	347	21.4	353	2 Q523G7_NOCRA	Q523G7 nocardia fa
35	339.5	20.9	341	2 Q79J80_CORGL	Q79J80 corynebacte
36	339.5	20.9	341	2 Q8NRL1_CORGL	Q8NRL1 corynebacte
37	329	20.3	347	2 Q523G8_NOCRA	Q523G8 nocardia fa
38	315.5	19.5	624	2 Q523G5_NOCRA	Q523G5 nocardia fa
39	311.5	19.2	338	2 Q6NBS3_CORBI	Q6NBS3 corynebacte
40	309.5	19.1	337	2 Q4JY22_CORJK	Q4JY22 corynebacte
41	297	18.3	360	2 Q8FLZ9_CORJK	Q8FLZ9 corynebacte
42	272.5	16.8	656	2 Q4JY21_CORJK	Q4JY21 corynebacte
43	267.5	16.5	126	2 P71483_MYCSC	P71483 mycobacteri
44	266	16.4	659	2 Q8FM00_CORJK	Q8FM00 corynebacte
45	265.5	16.4	319	2 Q522D5_NOCRA	Q522D5 nocardia fa

ALIGNMENTS

RESULT 1
MPT51_MYCBO STANDARD; PRT; 299 AA.
ID MPT51_MYCBO
AC POA4V7; Q31176; Q48923;
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE MPT51/MPB51 antigen precursor.
GN Name=mp51; Synonyms=fbpd, mpb51; Ordered locus names=MB3833c;
OS Mycobacterium bovis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium;
OC Mycobacterium tuberculosis complex.
OX NCBI_TaxID=1765;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=BCG /Tokyo/
RX MEDLINE=95242057; PubMed=7725062;
RA Ohara N., Kitaura H., Hotokezaka H., Nishiyama T., Wada N.,
RA Matsumoto S., Matsuo T., Naito M., Yamada T.,
RT "Characterization of the gene encoding the MPB51, one of the major
RT secreted protein antigens of Mycobacterium bovis BCG, and the
RT identification of the secreted protein closely related to the
RT fibronectin binding 85 complex."
RT Scand. J. Immunol. 41:433-442(1995).
RL [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=AF2122/97;
RX MEDLINE=2270107; PubMed=12788972; DOI=10.1073/pnas.1130426100;
RA Garner T., Biglmeier K., Camus J.-C., Medina N., Mansoor H.,
RA Payor M., Duthey S., Grondin S., Lacroix C., Monsemp C., Simon S.,
RA Harris B., Atkin R., Doggett J., Mayer R., Keating L., Wheeler P.R.,
RA Parthill J., Barrell B.G., Cole S.T., Gordon S.V., Hewinson R.G.,
RT "The complete genome sequence of Mycobacterium bovis."
RT Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882(2003).
RL -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: Belongs to the mycobacterial A85 antigen family.
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC
CC EMBL; D26486; BA05497.1; -; Genomic DNA.
CC EMBL; BX248347; CAD96019.1; -; Genomic DNA.
CC SMR; POA4V7; 33-299.
CC InterPro; IPR000801; Bacterase_put.
CC Pfam; PF00756; Bacterase; 1.
CC KAT Acyltransferase; Antigen; Complete proteome; Signal; Transferase.
CC STGNL 1 26
CC CHAIN 27 299 MPT51/MPB51 antigen.
CC FT
CC SQ SEQUENCE 299 AA; 31089 MW; 4E2R38F87ABD73E CRC64;

```

Query Match      99.7%; Score 1616; DB 1; Length 299;
Best Local Similarity 99.7%; Pred. No. 5.5e-104;
Matches 298; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MKGRSALLRLMTALISFGLGVAVAAPPTAKAPYENLWVPSMGRDIPVAFLAGGPH 60
DB 1 MKGRSALLRLMTALISFGLGVAVAAPPTAKAPYENLWVPSMGRDIPVAFLAGGPH 60
QY 61 AVYLLDAFNAGPVSVMWVTAGNMMNTLAGKISVAPAGAYSMYTNWEDGSKQMDTFL 120
DB 61 AVYLLDAFNAGPVSVMWVTAGNMMNTLAGKISVAPAGAYSMYTNWEDGSKQMDTFL 120
QY 121 SAELPWLAAVRGLAAGGHAAGAAGGAGAMALAAFHPRFGPAGSMGFLYPSNTTNN 180
DB 121 SAELPWLAAVRGLAAGGHAAGAAGGAGAMALAAFHPRFGPAGSMGFLYPSNTTNN 180
QY 181 GATTAAGGQOFGVDYTNMGAGAPOLGRKMHDPVHASLLAQNNTTRVWVMSPTNPGASDPA 240
DB 181 GATTAAGGQOFGVDYTNMGAGAPOLGRKMHDPVHASLLAQNNTTRVWVMSPTNPGASDPA 240
QY 241 AMIGQTAAAGNSRMFYNOYRSVGHNGHFPDPASGDNWGSWAPOLGAASGDIVGAIR 299
DB 241 AMIGQAALAMNSRMFTYNOYRSVGHNGHFPDPASGDNWGSWAPOLGAASGDIVGAIR 299

RESULT 2
MPT51 MYCTU      STANDARD;      PRT;      299 AA.
ID P04AV6, O33176, Q48923;
AC 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE MPT51/MPB51 antigen precursor.
GN Name=mp51; Synonyms=fbpd, mpb51; OrderedLocusNames=Rv3803c, MT3910;
ORFNames=MTY026.08c;
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium;
OC Mycobacterium tuberculosis complex.
OX NCBI_TaxId=1773;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=H37Rv;
RA Oettinger T., Andersen P.;
RT "Evidence for the secreted protein MPT51 from Mycobacterium tuberculosis is a T-cell antigen."
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=H37Rv;
MEDLINE=98295987; PubMed=9634230; DOI=10.1038/31159;
RA Cole S.T., Broesch R., Parkhill J., Garnier T., Churcher C.M.,
RA Harris D.E., Gordon S.V., Eiglmeier K., Gae S., Barry C.B. III,
RA Tekala F., Badcock K., Basham D., Brown D., Chillingworth T.,
RA Connor R., Davies R.M., Deavin K., Felwell T., Gentles S., Hamlin N.,
RA Holroyd S., Hornsby T., Jagsels K., Krogh A., McLean J., Moule S.,
RA Murphy L.D., Oliver S., Osborne J., Quail M.A., Rajandream M.A.,
RA Rogers J., Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence."
RL Nature 393:537-544(1998).
RN [3]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=CDC 1551 / Oshkosh;
MEDLINE=2206494; PubMed=12218036;
RX DOI=10.1126/JB.184.19.5479-5490.2002;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J.D., DeBoy R.T., Dodson R.J., Gwinn M.L., Haft D.H.,
RA Hickey E.K., Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D.,
RA Salzberg S.L., Delcher A., Ustebach T.R., Weidman J.F., Khouri H.M.,
RA Gill J., Mikula A., Bishai W., Jacobs W.R. Jr., Venter J.C.,
RA Fraser C.M.;

```

```

RT "Whole-genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains."
RL J. Bacteriol. 184:5479-5490(2002).
RN [4]
RP X-RAY CRYSTALLOGRAPHY (1.71 ANGSTROMS) OF 27-299.
RX PubMed=14672660; DOI=10.1016/j.jmb.2003.11.001;
RA Wilson R.A., Naughton W.N., Kremer L., Bees G.S., Pretreter K.;
RT "The structure of Mycobacterium tuberculosis MPT51 (FbpC1) defines a
RT new family of non-catalytic alpha/beta hydrolases."
RL J. Mol. Biol. 335:519-530(2004).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: Belongs to the mycobacterial A85 antigen family.
CC -----
CC This Swiss-Prot entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
DR EMBL; AJ002150; CA05211.1; -; Genomic DNA.
DR EMBL; BX842584; CA85642.1; -; Genomic DNA.
DR EMBL; AE000516; AAK48376.1; ALT_INIT; Genomic DNA.
DR PIR; G70887; G70887.
DR PDB; 1R88; X-ray; A/B=27-299.
DR SMR; P04AV6; 33-299.
DR TIGR; MT3910; -.
DR Tuberculist; Rv3803c; -.
DR InterPro; IPR000801; Esterase_put.
DR Pfam; PF00756; Esterase; 1.
DR 3D-structure; Acyltransferase; Antigen; Complete proteome; Signal;
KW Transferrase.
FT SIGNAL      1      26      Potential.
FT CHAIN       27     299     MPT51/MPB51 antigen.
FT FT          246     246     A -> T (in Ref. 1).
SQ SEQUENCE    299 AA; 31089 MW; 4E2E38F87AEDD73B CRC64;

Query Match      99.7%; Score 1616; DB 1; Length 299;
Best Local Similarity 99.7%; Pred. No. 5.5e-104;
Matches 298; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MKGRSALLRLMTALISFGLGVAVAAPPTAKAPYENLWVPSMGRDIPVAFLAGGPH 60
DB 1 MKGRSALLRLMTALISFGLGVAVAAPPTAKAPYENLWVPSMGRDIPVAFLAGGPH 60
QY 61 AVYLLDAFNAGPVSVMWVTAGNMMNTLAGKISVAPAGAYSMYTNWEDGSKQMDTFL 120
DB 61 AVYLLDAFNAGPVSVMWVTAGNMMNTLAGKISVAPAGAYSMYTNWEDGSKQMDTFL 120
QY 121 SAELPWLAAVRGLAAGGHAAGAAGGAGAMALAAFHPRFGPAGSMGFLYPSNTTNN 180
DB 121 SAELPWLAAVRGLAAGGHAAGAAGGAGAMALAAFHPRFGPAGSMGFLYPSNTTNN 180
QY 181 GATTAAGGQOFGVDYTNMGAGAPOLGRKMHDPVHASLLAQNNTTRVWVMSPTNPGASDPA 240
DB 181 GATTAAGGQOFGVDYTNMGAGAPOLGRKMHDPVHASLLAQNNTTRVWVMSPTNPGASDPA 240
QY 241 AMIGQTAAAGNSRMFYNOYRSVGHNGHFPDPASGDNWGSWAPOLGAASGDIVGAIR 299
DB 241 AMIGQAALAMNSRMFTYNOYRSVGHNGHFPDPASGDNWGSWAPOLGAASGDIVGAIR 299

RESULT 3
0745L2 MYCPA
ID 0745L2_MYCPA PRELIMINARY;      PRT;      300 AA.
AC 0745L2;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE FPBCL.
GN Name=fbpC1; OrderedLocusNames=MAF0217;
OS Mycobacterium paratuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium;

```

```

OC Mycobacterium avium complex (MAC).
OX NCBI_TaxID=1770;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=k10;
RA Li L., Bannantine J., Zhang Q., Amosin A., Alt D., Kapur V.;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB017227; AA002534.1; -; Genomic_DNA.
DR SMR; Q74512; 35-300.
DR InterPro; IPR000801; Esterase put.
DR InterPro; IPR001412; CRNA-synt_1.
DR Pfam; PF00756; Esterase; 1.
DR PROSITE; PS00178; AA_TRNA_Ligase_1; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 300 AA; 30929 MW; 11F27F49864423BD CRC64;

Query Match      82.0%; Score 1328.5; DB 2; Length 300;
Best Local Similarity 82.2%; Pred. No. 4,4e-84;
Matches 250; Conservative 16; Mismatches 29; Indels 9; Gaps 3;

QY 1 MKGRSALRLALMTALSFGLGVAVAAP-----TAKAPYENLWVPSMGKRDIPVAFL 55
DB 1 MRGVSLRLRFCLAVLAAGG---VALQPAAYGAAABAAGTESLWVSSAAMGRDIPVAFL 57
QY 56 AGGPAHVYLLDAFNAFVAGPVSNNWTAGNANTTLAAGKGISVVAAPAGAYSMTTNEQDGSQ 115
DB 58 AGGPAHVYLLDAFNAFVAGPVSNNWTAGNANTTLAAGKGISVVAAPAGAYSMTTNEQDGSQ 117
QY 116 WDTFSLASLPDMLAANRGLAPGHAAGVAAAGVAGMALAAPHDRGPGAGNSGFLYPS 175
DB 118 WDTFSLASLPDMLAANRGLAPGHAAGVAAAGVAGMALAAPHDRGPGAGNSGFLYPS 177
QY 176 NTTYNGAIAAGMOQFGVDVTNGMMGAPOLGRMKHDPVHVASLLAQNNTVWVWSPNPG 235
DB 178 STTYNGAIIAGLQOYGVVDNGMMGVOLGRMKHDPVHVASLLAQNNTVWVWSPNMG 237
QY 236 ASDPAMITGTYABAMGNSRMFTNYQYRSVGHNGHFDPPASGDNMGSGWAPOLGMSGDIV 295
DB 238 GDD-AAITGAGAGAGSSREFYQOYRSNGHGHFDPGSGDNMGSGWAPOLGMSGDIV 296
QY 296 GAIR 299
DB 297 GAIR 300

RESULT 4
MPT51_MYCLB STANDARD; PRT; 301 AA.
ID MPT51_MYCLB
AC 005868; Q50207;
DT 01-NOV-1995 (Rel. 32, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE MPT51 antigen precursor.
GN Name=mpc51; OrderedLocustNames=ML0098;
OS Mycobacterium leprae.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1769;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Thal53;
RA Yan Y.;
RT "Studies of MPT51 like protein of Mycobacterium leprae.",
RL Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.
[2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=TN;
RX MEDLINE=2118732; PubMed=11234002; DOI=10.1038/3505906;
RA Cole S.T., Bigmiller K., Parhill J., James K.D., Thomson N.R.,
RA Wheeler P.R., Honore N., Garnier T., Churcher C.M., Harris D.E.,
RA Mungall K.L., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R.M., Devlin K., Duthoy S., Feltham T., Fraser A., Hamlin N.,
RA Holroyd S., Hornsby T., Jagers K., Jorgensen C., Maclean J., Moule S.,

```

```

RA Murphy L.D., Oliver K., Quail M.A., Rajandream M.A., Rutherford K.M.,
RA Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R.,
RA Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,
RA Barrall B.G.;
RT "Massive gene decay in the leprosy bacillus."
RL Nature 409:1007-1011(2001).
RN [3]
RP NUCLEOTIDE SEQUENCE OF 1-220.
RX MEDLINE=93366419; PubMed=8359887;
RA Rinke de Wit T.F., Bekelle S., Oeland A., Wiles B., Janson A.A.M.,
RA Thole J.B.R.;
RT "The Mycobacterium leprae antigen 85 complex gene family:
RT identification of the genes for the 85A, 85C, and related MPT51
RT proteins."
RL Infect. Immun. 61:3642-3647(1993).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: Belongs to the mycobacterial A85 antigen family.
CC "This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed."
CC -----
DR EMBL; D50488; BAA09079.1; -; Genomic DNA.
DR EMBL; AL583917; CAC29606.1; -; Genomic DNA.
DR EMBL; Z21949; CAAT9947.1; -; Genomic DNA.
DR PIR; B86921; B86921.
DR PIR; S32111; S32111.
DR HSP; P31953; IDQZ.
DR SMR; Q05868; 36-301.
DR Leproma; ML0098; -.
DR InterPro; IPR000801; Esterase put.
DR Pfam; PF00756; Esterase; 1.
KW Acyltransferase; Antigen; Complete proteome; Signal; Transferase.
FT SIGNAL 1
FT CHAIN 37
FT CONFLICT 41
FT CONFLICT 286
FT FT
FT FT
SQ SEQUENCE 301 AA; 31191 MW; 3E666C8058FAB28 CRC64;

Query Match      77.2%; Score 1251; DB 1; Length 301;
Best Local Similarity 77.5%; Pred. No. 1e-78;
Matches 234; Conservative 23; Mismatches 41; Indels 4; Gaps 2;

QY 1 MKGRSALRLALMTALSFGLGVAVAAP--AKAPYENLWVPSMGKRDIPVAFLAG 57
DB 1 MRGLSAVVRVLCVAPALAVGFAAVALLAGTAGNAXKAGTESLWVPSNMGKRDIPVAFMAG 60
QY 58 GPHAVYLLDAFNAAGPVDVSNWVTAGNANTTLAAGKGISVVAAPAGAYSMTTWEQDGSQWD 117
DB 61 GPHAVYLLDAFNAALVDVSNWVTAGNANTTLAAGKGISVVAAPAGAYSMTTWEQDGSQWD 120
QY 118 TPLSASLPDMLAANRGLAPGHAAGVAAAGVAGMALAAPHDRPGAGSMGFLYPSNT 177
DB 121 TPLSASLPDMLAATKRGAPGHAAGVAAAGVAGMALAAPHDRPGAGSLGFLYPSST 180
QY 178 TTNGAIAAGMOQFGVDVTNGMMGAPOLGRMKHDPVHVASLLAQNNTVWVWSPNPGAS 237
DB 181 NTYNGAIIAGLQOYGVVDNGMMGAPOLGRMKHDPVHVASLLAQNNTVWVWSPNMG-G 239
QY 238 DPAAMITGTYABAMGNSRMFTNYQYRSVGHNGHFDPPASGDNMGSGWAPOLGMSGDIVGA 297
DB 240 DIDAMTGOAVASGSSREFYQOYRSVGHNGHFDPPGSGDNMGSGWAPOLGMSGDIVGA 299
QY 298 IR 299
DB 300 IR 301

```


OC Corynebacterineae; Mycobacteriaceae; Mycobacterium;
 OC Mycobacterium avium complex (MAC).
 OC NCBI_TaxID=1770;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=22481140; PubMed=12592709; DOI=10.1080/1042517021000019269;
 RA Dheendhavalan V., Shin K.S., Chang C.F., Chang C.D., Wang S.J.,
 RA McDonough S., McDonough P., Stehman S., Shin S., Torres A.,
 RA Chang Y.P.;
 RT Cloning and characterization of the genes coding for antigen 85A, 85B
 RT and 85C of Mycobacterium avium subsp. paratuberculosis.";
 RN DNA Seq. 13:287-294 (2002).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RA Chang Y.-P., Shin K.-S.;
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; A2780068; AAM73518.1; -; Genomic_DNA.
 DR HSPF; F31953; 1DQY.
 DR SMR; Q8K558; 49-326.
 DR InterPro; IPR000801; Beterase_put.
 DR Pfam; PF00756; Beterase; 1.
 SQ SEQUENCE 352 AA; 37754 MW; 39FPD0D7183975A8 CRC64;
 SO
 Query Match 35.1%; Score 569.5; DB 2; Length 352;
 Best Local Similarity 41.1%; Pred. No. 1.7e-31;
 Matches 130; Conservative 44; Mismatches 117; Indels 25; Gaps 9;
 QY 1 MKRSALT-RAIMTAALSFGLGVAVAA--EPTAKA-----APYENLWVSPSPMGDIP 51
 DB 10 LRGAALMPRLAIAGVAGSLSGVAVAGSPVAGSPKRGLEVELEVPSPSPMGNIK 69
 QY 52 VAFLAGPHAVYLLDAFNAGPDVSNWYTAGNMTTLAKGISVVAAPGAGVSMYTNBOD 111
 DB 70 VQGGGPHAVYLLDGLRADDDYNGMDINTPAPREYFQSGLSVIMPFQGGSPFSNNYQ 129
 QY 112 GS-----KQDFTLSAEIPDWLANRGLAPGHAAGVGAAGGYGMALAFHPDRFGF 164
 DB 130 SSGNGQYTYTKWETFLQEMPLMWQSNKQVSPAGNAVGLSMGSGSALIIAAYYPQGPY 189
 QY 165 AGSMGFLVPSNTTNGAIAAGMOQFGVDVNGMVGAPOLGRWKMHDPMVHASLIAQNT 224
 DB 190 AASLSEFLNPSBEGWMPFLIGLANMDSGGYNANSMWGPSTDPAMRRNDPMVQIPRVANNT 249
 QY 225 RVVWV-----SPTNAGASD-PAAMI-GQTAAEMGSRMFYNYQSVGSHGHFDPASGDN 278
 DB 250 RIWVYCGNGTPSDLGDNVPAKFLBGLT---LRINEQFQNNYAAAGRGVTFNPANGTH 306
 QY 279 GWGSAAPOLGAMSGDI 294
 DB 307 SWPYNQQLAMKPRDM 322
 RESULT 8
 A85C MYCBO STANDARD; PRT; 340 AA.
 ID A85C MYCBO P0A4V5; P31953; P96806;
 AC P0A4V5; P31953; P96806;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 10-MAY-2005 (Rel. 47, Last annotation update)
 DE Antigen 85-C precursor (85C) (Antigen 85 complex C) (Ag85C) (Mycoly1
 DE transferase 85C) (BC 2.3.1.-) (Fibronectin-binding protein C).
 GN Name=fbpc; Synonyms=mpc45; OrderedlocusNames=MB0134c;
 OS Mycobacterium bovis.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacterineae; Mycobacteriaceae; Mycobacterium;
 OC Mycobacterium tuberculosis complex.
 OC NCBI_TaxID=1765;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=AF2122/97;
 RX MEDLINE=2270107; PubMed=12788972; DOI=10.1073/pnas.1130426100;
 RA Garner T., Elgimeier K., Camus J.-C., Medina N., Mansoor H.,
 RA Pryor M., Duthey S., Grondin S., Lacroix C., Monsemp C., Simon S.,

RA Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,
 RA Parkhill J., Barrell B.G., Cole S.T., Gordon S.V., Hewinson R.G.;
 RT "The complete genome sequence of Mycobacterium bovis.";
 RL Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882 (2003).
 CC -1- FUNCTION: Protein of the antigen 85 complex are responsible for
 CC the high affinity of mycobacteria to fibronectin. Possesses a
 CC mycolyltransferase activity required for the biosynthesis of
 CC trehalose dimycolate (cord factor), a dominant structure necessary
 CC for maintaining cell wall integrity (By similarity).
 CC -1- SUBCELLULAR LOCATION: Secreted (By similarity).
 CC -1- SIMILARITY: Belongs to the mycobacterial A85 antigen family.
 CC -----
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC -----
 DR EMBL; BX248334; CAD92995.1; -; Genomic_DNA.
 DR SMR; P0A4V5; 49-328.
 DR InterPro; IPR000801; Beterase_put.
 DR Pfam; PF00756; Beterase; 1.
 KW Acyltransferase; Antigen; Complete proteome; Signal; Transferase.
 FT STGMAL 1 46 Potential.
 FT CHAIN 47 340 Antigen 85-C.
 FT REGION 102 112 Fibronectin-binding.
 FT ACT_SITE 170 170 Acyl-ester intermediate (By similarity).
 FT ACT_SITE 274 274 By similarity.
 FT ACT_SITE 306 306 By similarity.
 SQ SEQUENCE 340 AA; 36771 MW; 2D868C48D697988 CRC64;
 SO
 Query Match 34.8%; Score 564.5; DB 1; Length 340;
 Best Local Similarity 41.6%; Pred. No. 3.7e-31;
 Matches 131; Conservative 38; Mismatches 119; Indels 27; Gaps 8;
 QY 4 RGA---LRLALMTAA---LSFGL---GGVAVAEFTAKAAAYENLWVSPSPMGRIIPV 52
 DB 11 RSNATTLPRLLAIAGVAVLVYGLVETFGPATAVAFSRGDLPEYEQVPSASGRIKV 70
 QY 53 AFLAGPHAVYLLDAFNAGPDVSNWYTAGNMTTLAKGISVVAAPGAGVSMYTNBOD 112
 DB 71 QFGGGPHAVYLLDGLRADDDYNGMDINTPAPREYFQSGLSVIMPFQGGSPFSNNYQ 130
 QY 113 SK-----QMDFTLSAEIPDWLANRGLAPGHAAGVGAAGGYGMALAFHPDRFGFA 165
 DB 131 QSGNGQYTYTKWETFLREMPANWQANKGSPITGNAVGLSMGSGSALIIAAYYPQGPYA 190
 QY 166 GSNMGLVPSNTTNGAIAAGMOQFGVDVNGMVGAPOLGRWKMHDPMVHASLIAQNT 225
 DB 191 AASLSEFLNPSBEGWMPFLIGLANMDSGGYNANSMWGPSSDPAMRRNDPMVQIPRVANNT 250
 QY 226 VWVWV-----SPTNAGASD-PAAMI-GQTAAEMGSRMFYNYQSVGSHGHFDPASGDN 279
 DB 251 IWVYCGNGTPSDLGDNIPAKFLBGLT---LRINQFRTDYADGGNGVTFNPANGTH 307
 QY 280 WGSMAAPOLGAMSGDI 294
 DB 308 WPYNQQLAMKADI 322
 RESULT 9
 A85C MYCTU STANDARD; PRT; 340 AA.
 ID A85C MYCTU P0A4V6; P31953; P96806;
 AC P0A4V6; P31953; P96806;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 10-MAY-2005 (Rel. 47, Last annotation update)
 DE Antigen 85-C precursor (85C) (Antigen 85 complex C) (Ag85C) (Mycoly1
 DE transferase 85C) (BC 2.3.1.-) (Fibronectin-binding protein C).
 GN Name=fbpc; Synonyms=mpc45; OrderedlocusNames=Rv0129c; MT0137;
 OS Mycobacterium tuberculosis.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;

DB 191 ASLSGFLNBSGHWPTLIGLAMDSCGYNANSNMGSSDPAKKNDDPMVOIPLVANNTR 250
 QY 226 WWW-----SPTNPGASD-PAAMI-GOTABAMGNSRMFYNOYRSGVGHNGHDPASGSDNG 279
 DB 251 IWVYCGNGTSPDGLGDDNIPAKFTLEGLT---LRTNGTFRDPTAALADGNGNGVNFPPNGTHS 307
 QY 280 WGSWAPDLGAMSGDI 294
 DB 308 WPMWNEQLVAMKADI 322
 RESULT 10
 A85C_MYCLB STANDARD; PRT; 333 AA.
 ID A85C_MYCLB
 AC 005862;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 10-MAY-2005 (Rel. 47, Last annotation update)
 DE Antigen 85-C precursor (85C) (Antigen 85 complex C) (Ag85C) (Mycoly1
 DE transferase 85C) (BC 2.3.1.-) (Fibronectin-binding protein C).
 GN Name:fbpc; Synonyms=fbpc2; OrderedLocustNames=ML2655;
 OS Mycobacterium leprae.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxID=1769;
 RN NUCLEOTIDE SEQUENCE.
 RP MEDLINE=93366419; PubMed=8359887;
 RX Rinke de Wit T.F., Bekelle S., Orland A., Wiele B., Janson A.A.M.,
 RA Thole J.B.R.;
 RT "The Mycobacterium leprae antigen 85 complex gene family:
 RT Identification of the genes for the 85A, 85C, and related MPT51
 RT proteins.";
 RL Infect. Immun. 61:3642-3647(1993).
 RN [2]
 RA NUCLEOTIDE SEQUENCE.
 RP de Mendonca-Lima L.;
 RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=TV;
 RX MEDLINE=2118732; PubMed=11234002; DOI=10.1038/35059006;
 RA Cole S.T., Bigmiller K., Parkhill J., James K.D., Thomson N.R.,
 RA Wheeler P.R., Honore N., Garnier T., Churcher C.M., Harris D.B.,
 RA Mungall K.L., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Davies R.M., Devlin K., Duthoy S., Felwell T., Fraser A., Hamlin N.,
 RA Holroyd S., Hornsby T., Jagels K., Lacroix C., Maclean J., Moule S.,
 RA Murthy L.D., Oliver K., Quail M.A., Rajandream M.A., Rutherford K.M.,
 RA Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R.,
 RA Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,
 RA Barrall B.G.;
 RT "Massive gene decay in the leprosy bacillus".
 RL Nature 409:1007-1011(2001).
 CC -1- FUNCTION: Proteins of the antigen 85 complex are responsible for
 CC the high affinity of mycobacteria to fibronectin. Possesses a
 CC mycolyltransferase activity required for the biogenesis of
 CC trehalose dimycolate (cord factor), a dominant structure necessary
 CC for maintaining cell wall integrity (By similarity).
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- SIMILARITY: Belongs to the mycobacterial A85 antigen family.
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC EMBL; Z21951; CAAT9949.1; -; Genomic DNA.
 CC EMBL; M96649; AAA91865.1; -; Genomic DNA.
 CC EMBL; AL583926; CAC32187.1; -; Genomic DNA.
 CC PIR; S32114; S32114.
 CC HSP; P31953; IDOY.
 CC SWR; 005862; 49-328.

DR Leproma; ML2655; -.
 DR InterPro; IPR000801; Esterase_cat.
 DR Pfam; PF00756; Esterase_1.
 KW Acyltransferase; Antigen; Complete proteome; Signal; Transferase.
 FT SIGNAL 1 46 Potential.
 FT CHAIN 47 333 Antigen 85-C.
 FT REGION 102 112 Fibronectin-binding.
 FT ACT_SITE 170 170 Acyl-ester intermediate (By similarity).
 FT ACT_SITE 274 274 By similarity.
 FT ACT_SITE 306 306 By similarity.
 SQ SEQUENCE 333 AA; 36392 MW; 50EADF1F731EBEC2 CRC64;
 Query Match 33.8%; Score 548; DB 1; Length 333;
 Best local similarity 39.2%; Pred. No. 5e-30;
 Matches 118; Conservative 41; Mismatches 128; Indels 14; Gaps 5;
 QY 6 ALRALMIALSPGLGAVAAAEPTAKAAVEKMLVPSBPMGDIPIVAFLAGGHAYLL 65
 DB 24 AVIGTALLAGLVGVGTALAVAFSKRGLEVEYLVQVPSBPMGDIKIQFGGQGHAYLL 83
 QY 66 DAFNAGPDVSNWYTAGNANTLAGKISVVAAPAGAYSMYTNMEQ--DGSKQ-----WDT 118
 DB 84 DGLRQEDVNGMDINPAPFEEYHSGLSVIMPGGSSFYSNMYQPSQGNGHTYMET 143
 QY 119 FLGAEPLDWLANRGLAPGHAAGAAGGYGAMALAFHPRDRFAGSMGFLYPSNTT 178
 DB 144 FLTGEMPSYLGANKNVLPTGNAAVGLSMGSSALILASYPPQGFYPAASLSGFLNPSGW 203
 QY 179 TNGALIAAGMQFGCVDTNGKMGAPQLGRKWKHPVHASLLAQNTVYVW----SPTNP 234
 DB 204 WPTMIGLAMDSCGYNANSNMGSSDPAKKNDDPMVOIPLVANNTRIMWYCGNGANEL 263
 QY 235 GASD-PAAMI-GOTABAMGNSRMFYNOYRSGVGHNGHDPASGDNMGWAPDLGAMSGD 293
 DB 264 GGDNIAPKFL--ESLTISTHETIFONTYAAAGKNGVNFPPNGTHSPYMQDLVANKPD 321
 QY 294 I 294
 DB 322 I 322
 RESULT 11
 A85A_MYCBO STANDARD; PRT; 338 AA.
 ID A85A_MYCBO
 AC P044V3; P17944; P17996;
 DT 01-NOV-1990 (Rel. 16, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 10-MAY-2005 (Rel. 47, Last annotation update)
 DE Antigen 85-A precursor (85A) (Antigen 85 complex A) (Ag85A) (Mycoly1
 DE transferase 85A) (BC 2.3.1.-) (Fibronectin-binding protein A).
 GN Name:fbpa; Synonyms=mpc14; OrderedLocustNames=MB3834c;
 OS Mycobacterium bovis.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacterineae; Mycobacteriaceae; Mycobacterium;
 OC Mycobacterium tuberculosis complex.
 OX NCBI_TaxID=1765;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=BCG / Paris 1173 P2;
 RX MEDLINE=90326531; PubMed=2197602;
 RA de Wit L., de la Cuvelier A., Ooms J., Content J.;
 RT "Nucleotide sequence of the 32 kDa-protein gene (antigen 85 A) of
 RL Mycobacterium bovis BCG.";
 RL Nucleic Acids Res. 18:3995-3995(1990).
 RN [2]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=AF2122/97;
 RX MEDLINE=22709107; PubMed=12788972; DOI=10.1073/pnas.1130426100;
 RA Garnier T., Bigmiller K., Camus J.-C., Medina N., Mansoor H.,
 RA Payor M., Duthoy S., Grondin S., Lacroix C., Monnepe C., Simon S.,
 RA Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,
 RA Parkhill J., Barrall B.G., Cole S.T., Gordon S.V., Hewinson R.G.;
 RT "The complete genome sequence of Mycobacterium bovis.";

```

RL Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882(2003).
CC -1- FUNCTION: Proteins of the antigen 85 complex are responsible for
CC the high affinity of mycobacteria to fibronectin. Possesses a
CC mycolyltransferase activity required for the biogenesis of
CC trehalose dimycolate (cord factor), a dominant structure necessary
CC for maintaining cell wall integrity.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- PTM: Contains one disulfide bridge.
CC -1- SIMILARITY: Belongs to the mycobacterial A85 antigen family.
CC -----
CC This Swiss-Prot entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
DR EMBL: D26486; BAA05496.1; -; Genomic DNA.
DR EMBL: X53034; CAJ37206.1; -; Genomic DNA.
DR EMBL: BX248347; CAD96020.1; -; Genomic DNA.
DR PIR: S10326; S10326.
DR HSSP: P31952; 1F0N.
DR SMR: P0A4V3; 43-330.
DR InterPro: IPR000801; Esterase_put.
DR Pfam: PF00756; Esterase; 1.
KW Acyltransferase; Antigen; Complete proteome; Signal; Transferase.
FT CHAIN 1 42
FT REGION 43 338 Antigen 85-A.
FT ACT_SITE 101 111 Fibronectin-binding.
FT ACT_SITE 169 169 Acyl-ester intermediate (By similarity).
FT ACT_SITE 273 273 By similarity.
FT ACT_SITE 305 305 By similarity.
SQ SEQUENCE 338 AA; 35686 MW; 57B1CF95D07D52C0 CRC64;

Query Match 32.8%; Score 532; DB 1; Length 338;
Best Local Similarity 40.1%; Pred. No. 6.5e-29;
Matches 121; Conservative 37; Mismatches 126; Indels 18; Gaps 5;

QY 14 AALSFGI---GGVAVAEPTAKAPYENIMVPSKGRDIPVAFAGPH--AVYLLDA 67
DB 25 ALVSGLVGAVGTATATGAFSRPGLPEYIQLVPSKGRDITKQFQSGGANSPLATYLLDG 84
QY 68 FNAAGPVSNVVTAGNMMNTLAGIGISVYAPAGAYMYTWMEDQ-----SKQMDPL 120
DB 85 LRQDDPFGSGWDINTPAPFEMVDQGLSVMPVGVGSSFYSDWYPMACGKACQTYKKEFEL 144
QY 121 SAELPDMTLANRGLAGCGHAAVGAAGCGYGAAMALAAHPRPGPAGSGSEFLYPSMTTN 180
DB 145 TSSLPGLDANRRHVKPTGSAVVGLSMAASSALTATLYHPQGFYAGAMSGLLDPSQAMGP 204
QY 181 GATAAGCQOFGVDVTGMMGAPOLGRWKMDPWVHASILTAQNTRVWVW-----SPTNPGA 236
DB 205 TLIGLGMGADGGKAKADMDGPKRDPAMQRNDPLLVNCKLANTTRVWVVGNGKSPDLGG 264
QY 237 SDPAMNIQGTAAAGNSRMFYNQYRSVGGNGHFPDPASGDNCGSWAPOLGAMSGDIVG 296
DB 265 NNIPAKFLBGFVFTSNIK-FQDAVYNAAGGNGVDFDPDSTGHSWEYGAQLNAMKEDLQR 323
QY 297 AI 298
DB 324 AL 325

RESULT 12
ID A85A_MYCTU STANDARD; PRT; 338 AA.
AC P0A4V2; P17944; P17996;
DT 01-NOV-1990 (Rel. 16, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Antigen 85-A precursor (85A) (Antigen 85 complex A) (Ag85A) (Mycolyl
DE transferase 85A) (EC 2.3.1.1-) (Fibronectin-binding protein A).
GN Name=fbpA; Synonyms=mpc144; OrderedLocustNames=Rv3804c, MT3911;
ORFNames=MTV026.09c;

```

```

OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium;
OC Mycobacterium tuberculosis complex.
OC NCBI_TaxId=1773;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Erasmus;
RX MEDLINE=89379378; PubMed=2506131;
RA Bormans M., de Wit L., Volckaert G., Ooms J., de Bruyn J.,
RA Huygen K., van Vooren J.P., Stelandre M., Verhofstad R., Content J.,
RT "Cloning, sequence determination, and expression of a 32-kilodalton-
RT protein gene of Mycobacterium tuberculosis."
RL Infect. Immun. 57:3123-3130(1989).
RL [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Erasmus;
RX MEDLINE=9633337; PubMed=8757831;
RA Harth G., Lee B.Y., Wang J., Clemens D.L., Horwitz M.A.;
RT "Novel insights into the genetics, biochemistry, and
RT immunocytochemistry of the 30-kilodalton major extracellular protein
RT of Mycobacterium tuberculosis."
RL Infect. Immun. 64:3038-3047(1996).
RL [3]
RP ERRATUM.
RA Harth G., Lee B.Y., Wang J., Clemens D.L., Horwitz M.A.;
RL Infect. Immun. 65:852-852(1997).
RL [4]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=H37Rv;
RX MEDLINE=9825987; PubMed=9634230; DOI=10.1038/31159;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C.M.,
RA Harris D.E., Gordon S.V., Eigmeier K., Gas S., Barry C.E. III,
RA Tekala F., Badcock K., Basham D., Brown D., Chillingworth T.,
RA Connor R., Davies R.M., Devlin K., Feltwell T., Gentles S., Hamlin N.,
RA Holroyd S., Hornsby T., Jagels K., Krogh A., McLean M.A.,
RA Murphy L.D., Oliver S., Osborne J., Quail M.A., Rajandream M.A.,
RA Rogers J., Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sultson J.E., Taylor K., Whittam W.C., Umayam L.A., Ermolaeva M.D.,
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence."
RL Nature 393:537-544(1998).
RL [5]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=CDC 1551 / Oshkosh;
RX MEDLINE=22206494; PubMed=12218036;
RA DOI=10.1126/JB.184.19.5479-5490.2002;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J.D., Deboy R.T., Dodson R.J., Gwinn M.L., Haft D.H.,
RA Hickey B.K., Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D.,
RA Salzberg S.L., Delcher A., Utterback T.R., Weidman J.F., Kouri H.M.,
RA Gill J., Mink A., Bishai W., Jacobs W.R. Jr., Venter J.C.,
RA Frazer C.M.;
RT "Whole-genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains."
RL J. Bacteriol. 184:5479-5490(2002).
CC -1- FUNCTION: Proteins of the antigen 85 complex are responsible for
CC the high affinity of mycobacteria to fibronectin. Possesses a
CC mycolyltransferase activity required for the biogenesis of
CC trehalose dimycolate (cord factor), a dominant structure necessary
CC for maintaining cell wall integrity.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- PTM: Contains one disulfide bridge.
CC -1- SIMILARITY: Belongs to the mycobacterial A85 antigen family.
CC -----
CC This Swiss-Prot entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
DR EMBL: M27016; AAA50288.1; ALT_SEQ; Genomic DNA.
DR EMBL: U47335; AAC44295.1; -; Genomic DNA.

```

DR EMBL; BX842584; CAA17868.1; -; Genomic DNA.
 DR EMBL; AB000516; AAK48277.1; ALT_INIT; Genomic DNA.
 DR PIR; H70887; H70887.
 DR HSSP; P31952; 1F0N.
 DR SMR; POA4V2; 43-330.
 DR TIGR; MT3911; -.
 DR TubercuList; RV3804C; -.
 DR InterPro; IPR00801; Esterase_put.
 DR Pfam; PF00756; Esterase; 1.
 KW Acyltransferase; Antigen; Complete proteome; signal; transferase.
 FT SIGNAL 1 42
 FT CHAIN 43 338 Antigen 85-A.
 FT REGION 101 111 Fibrinectin-binding.
 FT ACT_SITE 169 169 Acyl-ester intermediate (By similarity).
 FT ACT_SITE 273 273 By similarity.
 FT ACT_SITE 305 305 By similarity.
 FT CONFLICT 24 26 GAA -> AR (in Ref. 1).
 FT CONFLICT 131 131 G -> R (in Ref. 1).
 FT CONFLICT 292 292 G -> R (in Ref. 1).
 FT CONFLICT 324 334 ALGATPNTGPA -> HWVPRPTGPR (in Ref. 1).
 FT SEQUENCE 338 AA; 35686 MM; 57B1CF95D07D52C0 CRC64;
 Query Match 32.8%; Score 532; DB 1; Length 338;
 Best Local Similarity 40.1%; Pred. No. 6.5e-29;
 Matches 121; Conservative 37; Mismatches 126; Indels 18; Gaps 5;
 QY 14 AALSFGI-----GGVAVAAPETAKAAYENIMVPSPSMGRIIPVAFIAGSPH--AVYILDA 67
 DB 25 AALVSGLVAGVGATATGAFSRPGLPVEYLQVPSPSMGRIIKVQFSGGANSPALYLLDG 84
 QY 68 FNAGPDVSNWVTAGNANTLAGKISVYAPAGAYSMYTNWEOG-----SKQMDTFL 120
 DB 85 LRQDDPSGMDINTPAFEWYDQGLSYVMFVGQSSFSYDWPYQACGKACQTYKMETFL 144
 QY 121 SAEIPDWLAAANGIAPGSHAAGVGAAGGAGMALAAFHDPFGAGSMGFLYPSNTTN 180
 DB 145 TSELPGILOANRHKYPFGSAVVGLSMAASSALTLAIYHPQGFYVAGAMSGILDPSQAMGP 204
 QY 181 GAIAAGMQPFGVDVTNGMGAPOLGRWKMDPWNHASLTAQNTTRVWVW-----SPTNPGA 236
 DB 205 TLIGLAMGDAGGYKASDMWPKEDPAWRDPLIANTRVWVYCGNGKPSDLGG 264
 QY 237 SDPAAMIGQTAEMANGSRMYNQYRSVGNHGHDPFASGDNGGSAAPOLGAMSGDIVG 296
 DB 265 NNLPAAKLEBGFVRTSNIK-FQDAVYAGGHNQVDFPDSGTHSEWYGAQLNANKPDLQR 323
 QY 297 AI 298
 DB 324 AL 325
 RESULT 13
 Q847N5_MYCTU PRELIMINARY; PRT; 338 AA.
 ID Q847N5_MYCTU PRELIMINARY; PRT; 338 AA.
 AC Q847N5;
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Secreted antigen Ag85A.
 OS Mycobacterium tuberculosis.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacteriaceae; Mycobacteriaceae; Mycobacterium;
 OC Mycobacterium tuberculosis complex.
 NC NCB1_TaxID=1773;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Lee C.F., Chang S.Y., Yu D.S.;
 RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY07395; AAC62004.1; -; Genomic DNA.
 DR HSSP; P31952; 1F0N.
 DR SMR; Q847N5; 43-330.
 DR InterPro; IPR00801; Esterase_put.
 DR Pfam; PF00756; Esterase; 1.

SQ SEQUENCE 338 AA; 35698 MM; 16B1CF95D07D486C CRC64;
 Query Match 32.8%; Score 532; DB 2; Length 338;
 Best Local Similarity 40.1%; Pred. No. 6.5e-29;
 Matches 121; Conservative 37; Mismatches 126; Indels 18; Gaps 5;
 QY 14 AALSFGI-----GGVAVAAPETAKAAYENIMVPSPSMGRIIPVAFIAGSPH--AVYILDA 67
 DB 25 AALVSGLVAGVGATATGAFSRPGLPVEYLQVPSPSMGRIIKVQFSGGANSPALYLLDG 84
 QY 68 FNAGPDVSNWVTAGNANTLAGKISVYAPAGAYSMYTNWEOG-----SKQMDTFL 120
 DB 85 LRQDDPSGMDINTPAFEWYDQGLSYVMFVGQSSFSYDWPYQACGKACQTYKMETFL 144
 QY 121 SAEIPDWLAAANGIAPGSHAAGVGAAGGAGMALAAFHDPFGAGSMGFLYPSNTTN 180
 DB 145 TSELPGILOANRHKYPFGSAVVGLSMAASSALTLAIYHPQGFYVAGAMSGILDPSQAMGP 204
 QY 181 GAIAAGMQPFGVDVTNGMGAPOLGRWKMDPWNHASLTAQNTTRVWVW-----SPTNPGA 236
 DB 205 TLIGLAMGDAGGYKASDMWPKEDPAWRDPLIANTRVWVYCGNGKPSDLGG 264
 QY 237 SDPAAMIGQTAEMANGSRMYNQYRSVGNHGHDPFASGDNGGSAAPOLGAMSGDIVG 296
 DB 265 NNLPAAKLEBGFVRTSNIK-FQDAVYAGGHNQVDFPDSGTHSEWYGAQLNANKPDLQR 323
 QY 297 AI 298
 DB 324 AL 325
 RESULT 14
 Q6PU6_MYCBO PRELIMINARY; PRT; 338 AA.
 ID Q6PU6_MYCBO PRELIMINARY; PRT; 338 AA.
 AC Q6PU6;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Secreted antigen 85A precursor.
 GN Name=hpa.
 OS Mycobacterium bovis BCG.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacteriaceae; Mycobacteriaceae; Mycobacterium;
 OC Mycobacterium tuberculosis complex.
 NC NCB1_TaxID=33892;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Liu F., Yu S., Chen Q., Nie J., Rui Y., Wu M., Zhu L., Li J.,
 RA Jiang X., Li Z.;
 RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY577803; AAS89650.1; -; Genomic DNA.
 DR SMR; Q6PU6; 43-330.
 DR InterPro; IPR00801; Esterase_put.
 DR Pfam; PF00756; Esterase; 1.
 KW Signal.
 FT SIGNAL 1 43 Potential.
 FT CHAIN 44 338 secreted antigen 85A.
 FT SEQUENCE 338 AA; 35686 MM; 57B1CF95D07D52C0 CRC64;
 Query Match 32.8%; Score 532; DB 2; Length 338;
 Best Local Similarity 40.1%; Pred. No. 6.5e-29;
 Matches 121; Conservative 37; Mismatches 126; Indels 18; Gaps 5;
 QY 14 AALSFGI-----GGVAVAAPETAKAAYENIMVPSPSMGRIIPVAFIAGSPH--AVYILDA 67
 DB 25 AALVSGLVAGVGATATGAFSRPGLPVEYLQVPSPSMGRIIKVQFSGGANSPALYLLDG 84
 QY 68 FNAGPDVSNWVTAGNANTLAGKISVYAPAGAYSMYTNWEOG-----SKQMDTFL 120
 DB 85 LRQDDPSGMDINTPAFEWYDQGLSYVMFVGQSSFSYDWPYQACGKACQTYKMETFL 144
 QY 121 SAEIPDWLAAANGIAPGSHAAGVGAAGGAGMALAAFHDPFGAGSMGFLYPSNTTN 180

```

DB      145 TSELPMQLANRHHVPTGSAVVGISMAASALTLATYHPQFPYAGAMSGILDPQAMGP 204
        ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
QY      181 GAIAGMOQFCGVDTNGMGAPOLGRWKHMDPWHASILAQNTTRVWVW----SPTNPGA 236
        :|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
DB      205 TLIGLAMDAGKXKASDMPKEDPAWQRNDPLINVKLIANNTRVWVYGNGKPSDLGG 264
        :|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
QY      237 SDPAAMIGQTAEAMGNSRMEFYNQYRSVGNHGHFDPASGDNMGSWAPOLGAMSGDIVG 296
        :|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
DB      265 NNLPAPFLSGFVATSNIK-FQDAVYNGGNGVGFDPDSTGTHSWEWGAQLNMKPDLOK 323
        :|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
QY      297 AI 298
        :|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
DB      324 AL 325
        :|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||

RESULT 15
A85A_MYCLB STANDARD; PRT; 330 AA.
ID A85A_MYCLB
AC 005861;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Antigen 85-A precursor (85A) (Antigen 85 complex A) (Ag85A) (Mycoly1
DE transferase 85A) (EC 2.3.1.-) (Fibronectin-binding protein A).
OS Name:fbpA; OrderedLocustNames=ML0097;
GN Mycobacterium leprae.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxId=1769;
NC 1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Tha153;
RC MEDLINE=95130993; PubMed=7829901;
RA yin Y.;
RT "Molecular cloning of alpha antigen like protein gene of Mycobacterium
RT leprae and its over production in Escherichia coli.";
RL Kaneshinogaku Zasshi 68:1330-1337(1994).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA de Mendonca-Lima L.;
RL Submitted (MAR-1996) to the EMBL/GenBank/DDBJ databases.
RN [3]
RC NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=TN;
RC MEDLINE=21128732; PubMed=11234002; DOI=10.1038/35059006;
RA Cole S.T., Bigmeier K., Parkhill J., James K.D., Thomson N.R.,
RA Wheeler P.R., Honore N., Garnier T., Churcher C.M., Harris D.B.,
RA Mungall K.L., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R.M., Devlin K., Duthoy S., Felwell T., Fraser A., Hamlin N.,
RA Holroyd S., Hornsby T., Jagsis K., Lacroix C., Maclean J., Moulle S.,
RA Murphy L.D., Oliver K., Quail M.A., Rajandream M.A., Rutherford K.M.,
RA Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R.,
RA Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,
RA Barrell B.G.;
RA "Massive gene decay in the leprosy bacillus.";
RL Nature 409:1007-1011(2001).
RN [4]
RP NUCLEOTIDE SEQUENCE OF 143-330.
RX MEDLINE=93366419; PubMed=835987;
RA Rinke de Wit T.F., Bekelle S., Oeland A., Wiles B., Janson A.A.M.,
RA Thole J.B.R.;
RT "The Mycobacterium leprae antigen 85 complex gene family:
RT identification of the genes for the 85A, 85C, and related MPT51
RT protein.";
RL Infect. Immun. 61:3642-3647(1993).
CC -1- FUNCTION: Proteins of the antigen 85 complex are responsible for
CC the high affinity of mycobacteria to fibronectin. Possesses a
CC mycolyltransferase activity required for the biogenesis of
CC trehalose dimycolate (cord factor), a dominant structure necessary
CC for maintaining cell wall integrity (By similarity).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- PTM: Contains one disulfide bridge (By similarity).

```

```

CC -1- SIMILARITY: Belongs to the mycobacterial A85 antigen family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; D43841; BAA07864.1; -; Genomic DNA.
CC EMBL; M90648; AAA31864.1; -; Genomic DNA.
CC EMBL; AL583917; CAC29605.1; -; Genomic DNA.
CC EMBL; Z218930; CAA79948.1; -; Genomic DNA.
CC PIR; A86921; A86921.
CC PIR; S32107; S32107.
CC HSSP; P31952; 1FON.
CC SMR; Q05861; 42-328.
CC LepToma; ML0097; -.
CC InterPro; IPR000801; Esterase_put.
CC Pfam; PF00756; Esterase; 1.
CC Acyltransferase; Antigen; Complete proteome; Signal; Transferase.
CC FT SIGNAL 1 42 Potential.
CC FT CHAIN 43 330 Antigen 85-A.
CC FT REGION 100 110 Fibronectin-binding.
CC FT ACT SITE 168 168 Acyl-ester intermediate (By similarity).
CC FT ACT SITE 272 272 By similarity.
CC FT ACT SITE 304 304 By similarity.
CC FT CONFLICT 149 149 O -> B (in Ref. 1).
CC FT SEQUENCE 330 AA; 35411 MW; 170C7C98C98FCSEC CRC64;

Query Match 32.1%; Score 521; DB 1; Length 330;
Best Local Similarity 38.9%; Pred. No. 3,7e-28;
Matches 119; Conservative 39; Mismatches 132; Indels 16; Gaps 5;

QY 4 RSALRLALMIALIS--FGAGVVAAPPTAKAPYENLWVPSPSMRDIPVALAGPH- 60
DB 16 RLRYVAMGVALLISALIGVGSAPAPAFSPGLPVLYLQVPSMSKRDIKVQNGGANS 75
QY 61 -AVYLLDAPNAGPDSVSNWYTAGNANTLAGKISVVAAPAGAVSMYTNNEQD----- 112
DB 76 PALYLLDGLRAQDDFGMDINTTAFEMYYQSGISVMPVGGQSSFYSDWYSPACGRAGCQ 135
QY 113 SKQMDFFLAELPDMILANRGLAPGHAAVGAAGGYGMALAAFFPDRRPFSGNSGFL 172
DB 136 TYKMETFLTSSELPQYLQSNKQIKPTGSAAVGSMAGLSALTALAYHPDPTVYGSNSGL 195
QY 173 YPSNTTNGAIIAAGMOQFCGVDTNGMGAPOLGRWKHMDPWHASILAQNTTRVWVW--- 229
DB 196 DPSNMGPSLLIGLAMDAGGYRAADWGSTPDAWGRANDPTVAVGTLIANTRIMVYCGN 255
QY 230 -SPTNPGASDPAAMIGQTAEAMGNSRMEFYNQYRSVGNHGHFDPASGDNMGSWAPOLG 288
DB 256 GKPTELGGNNLPKAKLLEGLVRTSNIK-FQGVYVAGGHNAAVFVFPDSTGTHSWEWGEQLN 314
QY 289 AMSGDI 294
DB 315 DMKCPDL 320

```

Search completed: April 14, 2006, 17:32:22
 Job time : 112.321 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: April 14, 2006, 17:32:59 ; Search time 27.6477 Seconds
(without alignments)
894.107 Million cell updates/sec

Title: US-10-620-246-42

Perfect score: 1621

Sequence: 1 MKGRSALLRLMTALSFGL.....WGSWAPQIGAMSGDIVCAIR 299

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA.*

1: /cgm2_6/ptodata/1/1aa/5.COMB.pep.*

2: /cgm2_6/ptodata/1/1aa/6.COMB.pep.*

3: /cgm2_6/ptodata/1/1aa/H.COMB.pep.*

4: /cgm2_6/ptodata/1/1aa/PCITUS.COMB.pep.*

5: /cgm2_6/ptodata/1/1aa/RS.COMB.pep.*

6: /cgm2_6/ptodata/1/1aa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1621	100.0	299	2	US-09-050-739-42
2	561.5	34.6	340	1	US-08-107-676-3
3	561.5	34.6	340	1	US-08-107-676-31
4	561.5	34.6	340	1	US-08-997-080-37
5	561.5	34.6	340	1	US-08-997-362-37
6	561.5	34.6	340	2	US-08-873-970-37
7	561.5	34.6	340	2	US-09-095-855-37
8	561.5	34.6	340	2	US-08-705-347A-37
9	561.5	34.6	340	2	US-09-324-542-37
10	561.5	34.6	340	2	US-09-205-426-37
11	561.5	34.6	340	2	US-09-200-643-37
12	561.5	34.6	340	2	US-09-295-820-3
13	561.5	34.6	340	2	US-09-295-820-31
14	561.5	34.6	308	2	US-08-508-761B-35
15	548	33.8	333	1	US-08-997-080-36
16	548	33.8	333	1	US-08-997-362-36
17	548	33.8	333	2	US-08-873-970-36
18	548	33.8	333	2	US-09-095-855-36
19	548	33.8	333	2	US-08-705-347A-36
20	548	33.8	333	2	US-09-324-542-36
21	548	33.8	333	2	US-09-205-426-36
22	548	33.8	333	2	US-09-200-643-36
23	532	32.8	321	2	US-08-508-761B-36
24	532	32.8	338	1	US-08-447-430A-39
25	532	32.8	338	1	US-08-107-676-28
26	532	32.8	338	1	US-08-997-080-32
27	532	32.8	338	1	US-08-997-080-34

28	532	32.8	338	1	US-08-997-362-32	Sequence 32, Appl
29	532	32.8	338	1	US-08-997-362-34	Sequence 34, Appl
30	532	32.8	338	2	US-08-873-970-32	Sequence 32, Appl
31	532	32.8	338	2	US-08-873-970-34	Sequence 34, Appl
32	532	32.8	338	2	US-09-095-855-32	Sequence 32, Appl
33	532	32.8	338	2	US-09-095-855-34	Sequence 34, Appl
34	532	32.8	338	2	US-08-705-347A-32	Sequence 32, Appl
35	532	32.8	338	2	US-08-705-347A-34	Sequence 34, Appl
36	532	32.8	338	2	US-09-324-542-32	Sequence 32, Appl
37	532	32.8	338	2	US-09-324-542-34	Sequence 34, Appl
38	532	32.8	338	2	US-09-205-426-32	Sequence 32, Appl
39	532	32.8	338	2	US-09-205-426-34	Sequence 34, Appl
40	532	32.8	338	2	US-09-200-643-32	Sequence 32, Appl
41	532	32.8	338	2	US-09-200-643-34	Sequence 34, Appl
42	532	32.8	338	2	US-09-342-673-39	Sequence 39, Appl
43	532	32.8	330	2	US-09-295-820-28	Sequence 28, Appl
44	523	32.3	330	1	US-08-997-080-30	Sequence 30, Appl
45	523	32.3	330	1	US-08-997-362-30	Sequence 30, Appl

ALIGNMENTS

RESULT 1
US-09-050-739-42
Sequence 42, Application US/09050739
Patent No. 6641814
GENERAL INFORMATION:
APPLICANT: ANDERSEN, Peter
APPLICANT: NIELSEN, Rikke
APPLICANT: OETTINGER, Thomas
APPLICANT: RASMUSSEN, Peter Black
APPLICANT: ROSENKRANDS, Ida
APPLICANT: WELDLING, Karin
APPLICANT: FLORIO, Walter
TITLE OF INVENTION: NUCLEIC ACIDS FRAGMENTS AND POLYPEPTIDE FRAGMENTS
TITLE OF INVENTION: DERIVED FROM M. TUBERCULOSIS
FILE REFERENCE: 670001-2002.1
CURRENT APPLICATION NUMBER: US/09/050,739
CURRENT FILING DATE: 1998-03-30
EARLIER APPLICATION NUMBER: 0376/97
EARLIER FILING DATE: 1997-04-02
EARLIER APPLICATION NUMBER: 1277/97
EARLIER FILING DATE: 1997-11-10
EARLIER APPLICATION NUMBER: 60/044,624
EARLIER FILING DATE: 1997-04-18
EARLIER APPLICATION NUMBER: 60/070,488
EARLIER FILING DATE: 1998-01-05
NUMBER OF SEQ ID NOS: 173
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 42
TYPE: PRT
ORGANISM: Mycobacterium tuberculosis
US-09-050-739-42
Query Match 100.0%; Score 1621; DB 2; Length 299;
Best Local Similarity 100.0%; Pred. No. 2.7e-136;
Matches 299; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MKGRSALLRLMTALSFGLGVAVAAEPKAKAPYENLWVSPSGMRDIPVAFIAGPH 60
DB 1 MKGRSALLRLMTALSFGLGVAVAAEPKAKAPYENLWVSPSGMRDIPVAFIAGPH 60
QY 61 AVYLLDAFNAGPVSNTVTGNAMNTIAGGISVYVAPGAGYSMTTWEDGSKQMTFL 120
DB 61 AVYLLDAFNAGPVSNTVTGNAMNTIAGGISVYVAPGAGYSMTTWEDGSKQMTFL 120
QY 121 SAEIPLDMLANRGAPGGAHVAAGGAGYAMLLAHPDRFGAGSMGFLVPSNTTNN 180
DB 121 SAEIPLDMLANRGAPGGAHVAAGGAGYAMLLAHPDRFGAGSMGFLVPSNTTNN 180
QY 181 GAIAAGQDPGGVDYTNMGWAPOLGRWKHDPMVHSLAQNNTTRVWWSPTNFGASDPA 240

Db 181 GATTAAGCGGCGVDTNGMAGAPOLGRKMKHDPTWVASLTAQNTTRVWVMSPTTPGASDPA 240
QY 241 AMIGOTAEAMGSRMFYNOYRSVGHNHGFDPASGNGMGWAPOLGAASGDIYGAI R 299
Db 241 AMIGOTAEAMGSRMFYNOYRSVGHNHGFDPASGNGMGWAPOLGAASGDIYGAI R 299

RESULT 2
US-08-107-676-3
; Sequence 3, Application US/08107676
; Patent No. 595356
; GENERAL INFORMATION:
; APPLICANT: Content, Jean
; APPLICANT: Dewitt, Lucas
; APPLICANT: Debruy, Jacqueline
; TITLE OF INVENTION: Mycobacterium polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding Them for Diagnosis and Control of
; TITLE OF INVENTION: Tuberculosis
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merchant & Gould
; STREET: 3100 No. 5955356west Center
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/107,676
; FILING DATE: 04-OCT-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/EP92/00268
; FILING DATE: 07-FEB-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP 91400388.4
; FILING DATE: 14-FEB-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Kowalchuk, Katherine M.
; REGISTRATION NUMBER: 36,848
; REFERENCE/DOCKET NUMBER: 8076.89USWO
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-332-9081
; TELEFAX: 612-332-5300
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 340 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-107-676-3

Query Match 34.6%; Score 561.5; DB 1; Length 340;
Best Local Similarity 41.3%; Pred. No. 6.4e-42;
Matches 130; Conservative 39; Mismatches 119; Indels 27; Gaps 8;

QY 4 RSA---LLRALMTIAA---LSFGL---GGVAVAAEPTAKAAYENILVPSMGRDIPV 52
Db 11 RSAAATTPRAVAIAAMAVLVYGVGFGPATGASRRGLPVRYIQVSSASGRITKV 70
QY 53 AFLAGGHAAYVLLDAFNAGPDVSNMTAGNANTLAKGISVAPAGAVSMYTNWEQDG 112
Db 71 QFGGGGHAAYVLLDGLRAQDDYNGMDINTPAFEERYQSGLSVIMPVGGSGSYTDWYQPS 130
QY 113 SK-----QMDPTLSAELDPMIAANGLAAPGHAAYGAAGCGGGAALAAFHDRGCPA 165
Db 131 QSNQNTYTYMKTFTLTREMPAMLOANKGVPTGNAAGVLSMGGSLIILAAYYPOQPYA 190

QY 166 GSNSEFLYPSNTTTNGAIAAGCQFGVDTNGMAGAPOLGRKMKHDPTWVASLTAQNTTR 225
Db 191 ASISGFLNPSSEGWPPTLIGLANMDSGYNANSMGSSDPAMRRNDPMVQIPRLVANNT R 250
QY 226 VWVW-----SPTNGASD-PAAMI-GOTAEAMGSRMFYNOYRSVGHNHGFDPASGNDG 279
Db 251 IWYICNGGTPSDIGDGNITPAKFLLEGLT---LRTNQTFRDTYAADGRAGVFNFPNGTIS 307

QY 280 WGSMAPOLGAMSGDI 294
Db 308 WPYWNEQLVAMKADI 322

RESULT 3
US-08-107-676-31
; Sequence 31, Application US/08107676
; Patent No. 595356
; GENERAL INFORMATION:
; APPLICANT: Content, Jean
; APPLICANT: Dewitt, Lucas
; APPLICANT: Debruy, Jacqueline
; TITLE OF INVENTION: Mycobacterium polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding Them for Diagnosis and Control of
; TITLE OF INVENTION: Tuberculosis
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merchant & Gould
; STREET: 3100 No. 5955356west Center
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/107,676
; FILING DATE: 04-OCT-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/EP92/00268
; FILING DATE: 07-FEB-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP 91400388.4
; FILING DATE: 14-FEB-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Kowalchuk, Katherine M.
; REGISTRATION NUMBER: 36,848
; REFERENCE/DOCKET NUMBER: 8076.89USWO
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-332-5300
; TELEFAX: 612-332-9081
; INFORMATION FOR SEQ ID NO: 31:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 340 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Mycobacterium tuberculosis
; IMMEDIATE SOURCE:
; CLONE: Protein sequence from antigen 85C from M.
; US-08-107-676-31

Query Match 34.6%; Score 561.5; DB 1; Length 340;
Best Local Similarity 41.3%; Pred. No. 6.4e-42;
Matches 130; Conservative 39; Mismatches 119; Indels 27; Gaps 8;

QY 4 RSA---LLRALMTIAA---LSFGL---GGVAVAAEPTAKAAYENILVPSMGRDIPV 52

DB 11 RSATTLPRVVAALAMGAVLVGVTGPGPATAGAFSRBGLPVEYIQQVPSASNGRDIKV 70
QY 53 AFLAGPHAVYLLDAFNAAGDVSNMVTAGNANMTLAGKGISVVAAGAAVSMYTNWEQDG 112
DB 71 QFQGGPHAVYLLDGLRAQDDYNGWDINTPAFEETYSGLSVIMPVGQSSFTTDWQPS 130
QY 113 SK-----QMDTFLSALPDMTLAANRGLAPGHAAGAAGGYGAMALAAFHDPREGFA 165
DB 131 OSNGQNTYTKMETFTLTREMPAMLOANKGVSPGTGNAAGLSMSGGSALITLAAYYQCPPYA 190
QY 166 GSMGFLYPSNTTNGAIAAGMOQFGVDNNGMVGAPQLGRKMKHDPWYASLLAQNTR 225
DB 191 ASLSGFLNPBSGWPPTLIGLANNDGGYNANSMWGSSDPAMKRNDDPMVOIPRLVANTR 250
QY 226 VWTW----SPTNPGASD-PAAMI-GOTABAMGSRMFYNOYRSVGNHGFDPASGDNG 279
DB 251 IMWYCGNGTSPSDGSDGDIIPAKFLEGLT---LRNQTFRDYAADGNGNGVFNFPNGTHS 307
QY 280 WGSWAPQLGMSGDI 294
DB 308 WPYWNEQLVAMKADI 322

RESULT 4

US-08-997-080-37
Sequence 37, Application US/08997080
Patent No. 596524
GENERAL INFORMATION:
APPLICANT: WATSON, JAMES D.
APPLICANT: TAN, PAUL U.J.
TITLE OF INVENTION: METHODS AND COMPOUNDS FOR THE TREATMENT OF IMMUNOLOGICALLY-
NUMBER OF SEQUENCES: 194
CORRESPONDENCE ADDRESS:
ADDRESSEE: Law Offices of Ann W. Speckman
STREET: 2601 Elliott Avenue, Suite 4185
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98121
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/997,080
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Sleath, Janet
REGISTRATION NUMBER: 37,007
REFERENCE/DOCKET NUMBER: 11000.1007
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-269-0565
TELEFAX: 206-269-0563
TELEX:
INFORMATION FOR SEQ ID NO: 37:
SEQUENCE CHARACTERISTICS:
LENGTH: 340 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-997-080-37

Query Match 34.6%; Score 561.5; DB 1; Length 340;
Best Local Similarity 41.3%; Pred. No. 6.4e-42;
Matches 130; Conservative 39; Mismatches 119; Indels 27; Gaps 8;
QY 4 RSA---LRLALMTAA---LSFGL---GGVAAAEPTAKAPYENLWPSPSNGRDIPV 52

DB 11 RSATTLPRVVAALAMGAVLVGVTGPGPATAGAFSRBGLPVEYIQQVPSASNGRDIKV 70
QY 53 AFLAGPHAVYLLDAFNAAGDVSNMVTAGNANMTLAGKGISVVAAGAAVSMYTNWEQDG 112
DB 71 QFQGGPHAVYLLDGLRAQDDYNGWDINTPAFEETYSGLSVIMPVGQSSFTTDWQPS 130
QY 113 SK-----QMDTFLSALPDMTLAANRGLAPGHAAGAAGGYGAMALAAFHDPREGFA 165
DB 131 OSNGQNTYTKMETFTLTREMPAMLOANKGVSPGTGNAAGLSMSGGSALITLAAYYQCPPYA 190
QY 166 GSMGFLYPSNTTNGAIAAGMOQFGVDNNGMVGAPQLGRKMKHDPWYASLLAQNTR 225
DB 191 ASLSGFLNPBSGWPPTLIGLANNDGGYNANSMWGSSDPAMKRNDDPMVOIPRLVANTR 250
QY 226 VWTW----SPTNPGASD-PAAMI-GOTABAMGSRMFYNOYRSVGNHGFDPASGDNG 279
DB 251 IMWYCGNGTSPSDGSDGDIIPAKFLEGLT---LRNQTFRDYAADGNGNGVFNFPNGTHS 307
QY 280 WGSWAPQLGMSGDI 294
DB 308 WPYWNEQLVAMKADI 322

RESULT 5

US-08-997-362-37
Sequence 37, Application US/08997362
Patent No. 5965287
GENERAL INFORMATION:
APPLICANT: Tan, Paul
APPLICANT: Hiyyama, Jun
APPLICANT: Visser, Elizabeth
APPLICANT: Skinner, Margot
APPLICANT: Scott, Linda
APPLICANT: Prestidge, Ross
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR
TREATMENT AND DIAGNOSIS OF MYCOBACTERIAL INFECTIONS
NUMBER OF SEQUENCES: 194
CORRESPONDENCE ADDRESS:
ADDRESSEE: Law Offices of Ann W. Speckman
STREET: 2601 Elliott Avenue, Suite 4185
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98121
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/997,362
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: U.S. Patent Application No. 5965287 08/873,970
FILING DATE: June 12, 1997
APPLICATION NUMBER: U.S. Patent Application No. 5965287 08/705,347
FILING DATE: August 29, 1996
ATTORNEY/AGENT INFORMATION:
NAME: Sleath, Janet
REGISTRATION NUMBER: 37,007
REFERENCE/DOCKET NUMBER: 11000.1002c2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-269-0565
TELEFAX: 206-269-0563
TELEX:
INFORMATION FOR SEQ ID NO: 37:
SEQUENCE CHARACTERISTICS:
LENGTH: 340 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear

/ MOLECULE TYPE: protein
US-08-997-362-37

Query Match 34.6%; Score 561.5; DB 1; Length 340;
Best Local Similarity 41.3%; Pred. No. 6,4e-42;
Matches 130; Conservative 39; Mismatches 119; Indels 27; Gaps 8;

QY 4 RSA---LRLALMIAA---LSFGL---GGVAVAEPTAKAAPEYENLWPSPMGRDIPV 52
DB 11 RSAATTLPRVALIAAAGAVLVGLVGFPGPATAGAFSRGLVEYVLQVPSASMGSDIKV 70
QY 53 AFLAGPPHAYVLLDAFNAGPDVSNWVTAGNAANTLAGKGISVVAAPAGAVSMYTNBQDG 112
DB 71 QFGGGPHAVYLLDGLRAODDYGMDINTPAFEERYQSGLSVIMPVGGQSSFTYDWYQPS 130
QY 113 SK-----QMDTFLSALPDMLAANRGLAPGHAAGAOGGYGAMALAAFFHPRGFA 165
DB 131 QSNQGVYTKMETFLTRFEMPAVLQANKVSPFGNAAVGLSMGSGSALILAAVYPOGFYPA 190
QY 166 GSNMGFLYPSNTTNGAIAAGMOQFGVDYTNMGWGAFLQGRWKMDPWWYASILLQNTNR 225
DB 191 ASLSGFLNPSGEGWPTLIGLANNDSGYNANSMWGPSSDPAMKRDPMVQIPRLVANTR 250
QY 226 VWTW----SPTNPGASD-PAAMI-GQTAAAMGNSRMFTQYNSVGGHNGHPTPPASGDNG 279
DB 251 IWYCGNGTSPSDGDNIPAKFLEGLT---LRTNQTFRDYAADGGRNGVFNFPNGTHS 307
QY 280 WGSMAPOLGMSGDI 294
DB 308 WPYWNEQLVAMKADI 322

RESULT 6

US-08-873-970-37

/ Sequence 37, Application US/08873970
/ Patent No. 6001361

/ GENERAL INFORMATION:

/ APPLICANT: Tan, Paul

/ APPLICANT: Hiyyama, Jun

/ APPLICANT: Vlieser, Elizabeth

/ APPLICANT: Skinner, Margaret

/ APPLICANT: Scott, Linda

/ APPLICANT: Prestidge, Rose

/ TITLE OF INVENTION: COMPOUNDS AND METHODS FOR

/ TITL OF INVENTION: TREATMENT AND DIAGNOSIS OF MYCOBACTERIAL INFECTIONS

/ NUMBER OF SEQUENCES: 106

/ CORRESPONDENCE ADDRESS:

/ ADDRESSER: Law Offices of Ann W. Speckman

/ STREET: 2601 Elliott Avenue, Suite 4185

/ CITY: Seattle

/ STATE: WA

/ COUNTRY: USA

/ ZIP: 98121

/ COMPUTER READABLE FORM:

/ MEDIUM TYPE: Diskette

/ COMPUTER: IBM Compatible

/ OPERATING SYSTEM: DOS

/ SOFTWARE: FastSeq for Windows Version 2.0

/ CURRENT APPLICATION DATA:

/ APPLICATION NUMBER: US/08/873,970

/ FILING DATE:

/ CLASSIFICATION: 435

/ PRIOR APPLICATION DATA:

/ APPLICATION NUMBER: 08/705,347

/ FILING DATE: 29-AUG-1996

/ ATTORNEY/AGENT INFORMATION:

/ NAME: Sleath, Janet

/ REGISTRATION NUMBER: 37,007

/ TELECOMMUNICATION INFORMATION:

/ TELEPHONE: 206-269-0565

/ TELEFAX: 206-269-0563

/ TELEX:

/ INFORMATION FOR SEQ ID NO: 37:

/ SEQUENCE CHARACTERISTICS:

/ LENGTH: 340 amino acids

/ TYPE: amino acid

/ STRANDEDNESS: single

/ TOPOLOGY: linear

/ MOLECULE TYPE: protein

/ US-08-873-970-37

Query Match 34.6%; Score 561.5; DB 2; Length 340;
Best Local Similarity 41.3%; Pred. No. 6,4e-42;
Matches 130; Conservative 39; Mismatches 119; Indels 27; Gaps 8;

QY 4 RSA---LRLALMIAA---LSFGL---GGVAVAEPTAKAAPEYENLWPSPMGRDIPV 52
DB 11 RSAATTLPRVALIAAAGAVLVGLVGFPGPATAGAFSRGLVEYVLQVPSASMGSDIKV 70
QY 53 AFLAGPPHAYVLLDAFNAGPDVSNWVTAGNAANTLAGKGISVVAAPAGAVSMYTNBQDG 112
DB 71 QFGGGPHAVYLLDGLRAODDYGMDINTPAFEERYQSGLSVIMPVGGQSSFTYDWYQPS 130
QY 113 SK-----QMDTFLSALPDMLAANRGLAPGHAAGAOGGYGAMALAAFFHPRGFA 165
DB 131 QSNQGVYTKMETFLTRFEMPAVLQANKVSPFGNAAVGLSMGSGSALILAAVYPOGFYPA 190
QY 166 GSNMGFLYPSNTTNGAIAAGMOQFGVDYTNMGWGAFLQGRWKMDPWWYASILLQNTNR 225
DB 191 ASLSGFLNPSGEGWPTLIGLANNDSGYNANSMWGPSSDPAMKRDPMVQIPRLVANTR 250
QY 226 VWTW----SPTNPGASD-PAAMI-GQTAAAMGNSRMFTQYNSVGGHNGHPTPPASGDNG 279
DB 251 IWYCGNGTSPSDGDNIPAKFLEGLT---LRTNQTFRDYAADGGRNGVFNFPNGTHS 307
QY 280 WGSMAPOLGMSGDI 294
DB 308 WPYWNEQLVAMKADI 322

RESULT 7

US-09-095-855-37

/ Sequence 37, Application US/09095855
/ Patent No. 6160093

/ GENERAL INFORMATION:

/ APPLICANT: Tan, Paul

/ APPLICANT: Vlieser, Elizabeth

/ APPLICANT: Skinner, Margaret

/ APPLICANT: Prestidge, Rose

/ TITLE OF INVENTION: Compounds and Methods for

/ TITL OF INVENTION: Treatment and Diagnosis of Mycobacterial Infections

/ NUMBER OF SEQUENCES: 208

/ CORRESPONDENCE ADDRESS:

/ ADDRESSER: Law Offices of Ann W. Speckman

/ STREET: 2601 Elliott Avenue, Suite 4185

/ CITY: Seattle

/ STATE: WA

/ COUNTRY: USA

/ ZIP: 98121

/ COMPUTER READABLE FORM:

/ MEDIUM TYPE: Diskette

/ COMPUTER: IBM Compatible

/ OPERATING SYSTEM: DOS

/ SOFTWARE: FastSeq for Windows Version 2.0

/ CURRENT APPLICATION DATA:

/ APPLICATION NUMBER: US/09/095,855

/ FILING DATE:

/ CLASSIFICATION:

/ PRIOR APPLICATION DATA:

/ APPLICATION NUMBER: 08/705,347

/ FILING DATE: 29-AUG-1996

/ APPLICATION NUMBER: 08/873,970

/ FILING DATE: 12-JUN-1997

/ APPLICATION NUMBER: 08/997,362

/ FILING DATE: 23-DEC-1997

ATTORNEY/AGENT INFORMATION:
NAME: Sleath, Janet
REGISTRATION NUMBER: 37,007
REFERENCE/DOCKET NUMBER: 11000.1002c3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-269-0565
TELEFAX: 206-269-0563
TELEX:
INFORMATION FOR SEQ ID NO: 37:
SEQUENCE CHARACTERISTICS:
LENGTH: 340 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULAR TYPE: protein
US-09-095-855-37

Query Match 34.6%; Score 561.5; DB 2; Length 340;
Best Local Similarity 41.3%; Pred. No. 6.4e-42;
Matches 130; Conservative 39; Mismatches 119; Indels 27; Gaps 8;

4 RSA---LRLALMTA---LSFGL---GGVAVAABPTAKAAYENIMVPSMGRDIPV 52
11 RSAATLPRVAALAMGAVLVGLVGTGFPATAGAFSRGLPEVYLQVPSASMKRDIKV 70

53 APLAGPHAYVLLDAFNAGPDVSNMVTAGNMTLAGKGISVVAAPAGAAVSMYTNMEQDG 112
71 QFGGPHAYVLLDGLRAODDYNMGDINTAFBEYYSGLSVIMPVGGSSFTTDWQPS 130

113 SK-----QMDTFLSALPDWLANRGLAPGHAAGAAGGYGAMALAAFHDRGFA 165
131 QSNQNTYTKMETFLTEPMAMLOANKGVSPGNAAVGLSMGSSALILAAVYPOQFPYA 190

166 GMSGFLYPSNTTNGAIAAGMOQFGVDNMGWGAQLGRKMDPWYASLLAONTR 225
191 ASLSGFLNPSBGWPTLIGLAMDSSGYNANSMGSSDPAMKRNDDPVQIPRLVANTR 250

226 VMTW----SPTNPGASD-PAAMI-GOTAAAGNSRMFYNOYRSVGHNGHFDPPASGNG 279
251 IWVYCGNGTSDLDGNDIPAKFLEGLT---LRNQTFRDYAADGNGNGVFNFPNGTHS 307

280 WGSNAPQLGAMSGDI 294
308 WPYWNEQLVAMKADI 322

Db

RESULT 8
US-08-705-347A-37
Sequence 37, Application US/08705347A
Patent No. 6284255
GENERAL INFORMATION:
APPLICANT: Tan, Paul
APPLICANT: Hiyaama, Jun
APPLICANT: Visser, Elizabeth
APPLICANT: Skinner, Margot
APPLICANT: Scott, Linda
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT AND
TITLE OF INVENTION: DIAGNOSIS OF MYCOBACTERIAL INFECTIONS
NUMBER OF SEQUENCES: 55
CORRESPONDENCE ADDRESS:
ADDRESSER: Speckman Plicard PLLC
STREET: 2601 Elliott Avenue, Suite 4185
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98121
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/705,347A

FILING DATE: 28-AUG-1996
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Sleath, Janet
REGISTRATION NUMBER: 37,007
REFERENCE/DOCKET NUMBER: 11000.1002
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206.269.0563
TELEFAX: 206.269.0563
INFORMATION FOR SEQ ID NO: 37:
SEQUENCE CHARACTERISTICS:
LENGTH: 340 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULAR TYPE: protein
US-08-705-347A-37

Query Match 34.6%; Score 561.5; DB 2; Length 340;
Best Local Similarity 41.3%; Pred. No. 6.4e-42;
Matches 130; Conservative 39; Mismatches 119; Indels 27; Gaps 8;

4 RSA---LRLALMTA---LSFGL---GGVAVAABPTAKAAYENIMVPSMGRDIPV 52
11 RSAATLPRVAALAMGAVLVGLVGTGFPATAGAFSRGLPEVYLQVPSASMKRDIKV 70

53 APLAGPHAYVLLDAFNAGPDVSNMVTAGNMTLAGKGISVVAAPAGAAVSMYTNMEQDG 112
71 QFGGPHAYVLLDGLRAODDYNMGDINTAFBEYYSGLSVIMPVGGSSFTTDWQPS 130

113 SK-----QMDTFLSALPDWLANRGLAPGHAAGAAGGYGAMALAAFHDRGFA 165
131 QSNQNTYTKMETFLTEPMAMLOANKGVSPGNAAVGLSMGSSALILAAVYPOQFPYA 190

166 GMSGFLYPSNTTNGAIAAGMOQFGVDNMGWGAQLGRKMDPWYASLLAONTR 225
191 ASLSGFLNPSBGWPTLIGLAMDSSGYNANSMGSSDPAMKRNDDPVQIPRLVANTR 250

226 VMTW----SPTNPGASD-PAAMI-GOTAAAGNSRMFYNOYRSVGHNGHFDPPASGNG 279
251 IWVYCGNGTSDLDGNDIPAKFLEGLT---LRNQTFRDYAADGNGNGVFNFPNGTHS 307

280 WGSNAPQLGAMSGDI 294
308 WPYWNEQLVAMKADI 322

Db

RESULT 9
US-09-324-542-37
Sequence 37, Application US/09324542
Patent No. 6328978
GENERAL INFORMATION:
APPLICANT: Watson, James D.
APPLICANT: Tan, Paul L.J.
APPLICANT: Prestidge, Ross
APPLICANT: Methods and Compounds for the Treatment
TITLE OF INVENTION: of Immunologically-Mediated Skin Disorders
FILE REFERENCE: 11000.1007c1
CURRENT APPLICATION NUMBER: US/09/324,542
CURRENT FILING DATE: 1999-06-02
EARLIER APPLICATION NUMBER: US 08/997,080
EARLIER FILING DATE: 1997-12-23
NUMBER OF SEQ ID NOS: 194
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 37
LENGTH: 340
TYPE: PRT
ORGANISM: Mycobacterium tuberculosis
US-09-324-542-37

Query Match 34.6%; Score 561.5; DB 2; Length 340;
Best Local Similarity 41.3%; Pred. No. 6.4e-42;
Matches 130; Conservative 39; Mismatches 119; Indels 27; Gaps 8;

```
QY 4 RSA---LRLALMTAA---LSFGL---GGVAAAEPTAKAAYENLMPSPSGRDIPIV 52
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 11 RSAATTLPRRVAIAAMGAVLVGTVFGPATAGASRPGLPVEYLOVPSASMGRIKIV 70
QY 53 AFLAGGPHAVYLLDAFNAGPDVSNWVTAGNANMTLAKGISVAAPAGAVSMITTNEODG 112
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 71 QFQGGGPHAVYLLDGLRAODDYNMGDINTPAFEBYVQSGLSVIMPVGQSSPYTDWYQPS 130
QY 113 SK-----QMDTFLSAELPDMTLAANGGLAPGGHAAVGAAGGAGMALAAPHDRFGFA 165
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 131 QSNQONTYTKMETFLTRKMPAMLOANKGVSPGTGNAAGLSMSGGSLILTAAYTPQCPYA 190
QY 166 GSNMGFLYPSNTTNGAIAAGMOQFGGVDNMGWGAPOLGKRWKMDPWHASLLAQNTR 225
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 191 ASISGFLNPSEGWMPTLIGLANMDSGGYNANSMWGPSSDPAMKRNDEPMVOIPRLVANTR 250
QY 226 VWWV---SPTNPGASD-PAAMI-GQTAEAMGNSRMFYNOYRSVGGHGHFDPASGDNG 279
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 251 IWYICGNGTPSDIGSDNIPAKFLEGTL---LRTNQTFRDTYAADGGRNGVFNFPNGTHS 307
QY 280 WGSMAPOLGAMSGDI 294
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 308 WPYMNEQLVAMKADI 322
```

RESULT 10

```
US-09-205-426-37
; Sequence 37, Application US/09205426
; Patent No. 6406704
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; TITLE OF INVENTION: Compounds and Methods for Treatment and
; TITLE OF INVENTION: Diagnosis of Mycobacterial Infections
; FILE REFERENCE: 11000.1002c4
; CURRENT APPLICATION NUMBER: US/09/205.426
; EARLIER FILING DATE: 1998-12-04
; EARLIER APPLICATION NUMBER: 09/095,855
; EARLIER FILING DATE: 1998-06-11
; EARLIER APPLICATION NUMBER: 08/997,362
; EARLIER FILING DATE: 1997-12-23
; EARLIER APPLICATION NUMBER: 08/873,970
; EARLIER FILING DATE: 1997-06-12
; EARLIER APPLICATION NUMBER: 08/705,347
; EARLIER FILING DATE: 1996-08-29
; NUMBER OF SEQ ID NOS: 208
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 37
; LENGTH: 340
; TYPE: PRF
; ORGANISM: Mycobacterium tuberculosis
US-09-205-426-37
```

```
Query Match 34.6%; Score 561.5; DB 2; Length 340;
Best Local Similarity 41.3%; Pred. No. 6,4e-42;
Matches 130; Conservative 39; Mismatches 119; Indels 27; Gaps 8;
```

```
QY 4 RSA---LRLALMTAA---LSFGL---GGVAAAEPTAKAAYENLMPSPSGRDIPIV 52
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 11 RSAATTLPRRVAIAAMGAVLVGTVFGPATAGASRPGLPVEYLOVPSASMGRIKIV 70
QY 53 AFLAGGPHAVYLLDAFNAGPDVSNWVTAGNANMTLAKGISVAAPAGAVSMITTNEODG 112
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 71 QFQGGGPHAVYLLDGLRAODDYNMGDINTPAFEBYVQSGLSVIMPVGQSSPYTDWYQPS 130
QY 113 SK-----QMDTFLSAELPDMTLAANGGLAPGGHAAVGAAGGAGMALAAPHDRFGFA 165
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 131 QSNQONTYTKMETFLTRKMPAMLOANKGVSPGTGNAAGLSMSGGSLILTAAYTPQCPYA 190
QY 166 GSNMGFLYPSNTTNGAIAAGMOQFGGVDNMGWGAPOLGKRWKMDPWHASLLAQNTR 225
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 191 ASISGFLNPSEGWMPTLIGLANMDSGGYNANSMWGPSSDPAMKRNDEPMVOIPRLVANTR 250
```

```
QY 226 VWWV---SPTNPGASD-PAAMI-GQTAEAMGNSRMFYNOYRSVGGHGHFDPASGDNG 279
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 251 IWYICGNGTPSDIGSDNIPAKFLEGTL---LRTNQTFRDTYAADGGRNGVFNFPNGTHS 307
QY 280 WGSMAPOLGAMSGDI 294
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 308 WPYMNEQLVAMKADI 322
```

RESULT 11

```
US-09-200-643-37
; Sequence 37, Application US/09200643
; Patent No. 6410720
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; TITLE OF INVENTION: Compounds and Methods for Treatment and
; TITLE OF INVENTION: Diagnosis of Mycobacterial Infections
; FILE REFERENCE: 11000.1002CON
; CURRENT APPLICATION NUMBER: US/09/200.643
; EARLIER FILING DATE: 1998-11-05
; EARLIER APPLICATION NUMBER: 08/705,347
; EARLIER FILING DATE: 1996-08-29
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 37
; LENGTH: 340
; TYPE: PRF
; ORGANISM: Mycobacterium tuberculosis
US-09-200-643-37
```

```
Query Match 34.6%; Score 561.5; DB 2; Length 340;
Best Local Similarity 41.3%; Pred. No. 6,4e-42;
Matches 130; Conservative 39; Mismatches 119; Indels 27; Gaps 8;
```

```
QY 4 RSA---LRLALMTAA---LSFGL---GGVAAAEPTAKAAYENLMPSPSGRDIPIV 52
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 11 RSAATTLPRRVAIAAMGAVLVGTVFGPATAGASRPGLPVEYLOVPSASMGRIKIV 70
QY 53 AFLAGGPHAVYLLDAFNAGPDVSNWVTAGNANMTLAKGISVAAPAGAVSMITTNEODG 112
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 71 QFQGGGPHAVYLLDGLRAODDYNMGDINTPAFEBYVQSGLSVIMPVGQSSPYTDWYQPS 130
QY 113 SK-----QMDTFLSAELPDMTLAANGGLAPGGHAAVGAAGGAGMALAAPHDRFGFA 165
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 131 QSNQONTYTKMETFLTRKMPAMLOANKGVSPGTGNAAGLSMSGGSLILTAAYTPQCPYA 190
QY 166 GSNMGFLYPSNTTNGAIAAGMOQFGGVDNMGWGAPOLGKRWKMDPWHASLLAQNTR 225
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 191 ASISGFLNPSEGWMPTLIGLANMDSGGYNANSMWGPSSDPAMKRNDEPMVOIPRLVANTR 250
QY 226 VWWV---SPTNPGASD-PAAMI-GQTAEAMGNSRMFYNOYRSVGGHGHFDPASGDNG 279
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 251 IWYICGNGTPSDIGSDNIPAKFLEGTL---LRTNQTFRDTYAADGGRNGVFNFPNGTHS 307
QY 280 WGSMAPOLGAMSGDI 294
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 308 WPYMNEQLVAMKADI 322
```

RESULT 12

```
US-09-295-820-3
; Sequence 3, Application US/09295820
; Patent No. 6638511
; GENERAL INFORMATION:
```

```
APPLICANT: Content, Jean
Dewitt, Lucas
DeBruyn, Jacqueline
TITLE OF INVENTION: Mycobacterium Polypeptides and Nucleic
Acids Encoding Them for Diagnosis and Control of
Tuberculosis
NUMBER OF SEQUENCES: 38
```


Db 308 WPYMNQOLVAMKADI 322

RESULT 14

US-08-508-761B-35

Sequence 35, Application US/08508761B

Patent No. 6027920

GENERAL INFORMATION:

APPLICANT: Joliff, Gwennael

APPLICANT: Guyonvarch, Armel

APPLICANT: Purification, Relano

APPLICANT: Duchillon, Francis

APPLICANT: Renaud, Michel

TITLE OF INVENTION: System for Protein Expression and

NUMBER OF SEQUENCES: 37

CORRESPONDENCE ADDRESS: 37

ADDRESSER: Jacobson, Price, Holman & Stern, PLLC

STREET: 400 Seventh St. N.W.

CITY: Washington D.C.

COUNTRY: U.S.A.

ZIP: 20004

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/508,761B

FILING DATE: 31-JUL-1995

CLASSIFICATION:

PRIOR APPLICATION DATA: FR 91/09652

FILING DATE: 29-JUL-1991

PRIOR APPLICATION DATA: FR 91/09870

FILING DATE: 02-AUG-1991

ATTORNEY/AGENT INFORMATION:

NAME: Player, William E.

REGISTRATION NUMBER: 31,409

REFERENCE/DOCKET NUMBER: P58525NA

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 638-6666

TELEFAX: (202) 393-5350

INFORMATION FOR SEQ ID NO: 35:

SEQUENCE CHARACTERISTICS:

LENGTH: 308 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

MOLECULE TYPE: protein

ORIGINAL SOURCE:

ORGANISM: M. tuberculosis

US-08-508-761B-35

Query Match 34.6%; Score 561; DB 2; Length 308;

Best Local Similarity 41.3%; Pred. No. 6,3e-42;

Matches 123; Conservative 39; Mismatches 117; Indels 20; Gaps 6;

Db 14 AALSFGI-----GSVAVAEPTAKAAPYENLMVPSPSMGRIPIVAPFLAGSHAYILLDAFN 69

Db 6 AVLVYGLVGTFFGPATAGAPSRPGLPYEYLQVPSASMGRIKIQVPOGGPHAYILLDGLR 65

Db 70 AGPDVSNWVTAGNANMTLAGKGISVYVAPAGAVSMYTMEDQSK-----QMDFTLSA 122

Db 66 AADDYNGMDINTPAFEERYQSGLSVIMPVGGQSSFTTDMYQPSQSNQNTTYKMETFLTR 125

Db 123 ELDPMLAANRGLAPGHAIVAGAAQGGYGAMALAAFHEDRFAGSGNSGFLYPSNTTNGA 182

Db 126 EWPAAVLQANKGVSPGTGAIVAGLSWSGSSALIIAAVYPOQPYAASLSGFLNBSGWWPTL 185

Db 183 IAAAGMOQFGGVDTNGMGAPQLGRWKMDPVTASLLAQNTRVYMW-----SPNPGASD 238

Db 186 IGLANDSGGYNANSMWGPSSDPAWKNDPMVOI PRLVANNTRIMWYCGNGTSPSDLDGDN 245

QY 239 -PAAMI-GQTAAEMNSRMFTYQYRSVGHNHGFDPASGDNGSWAPQLGAMSDI 294
Db 246 IPAKFLFGLT--LRNQTFRDITYADGGRNGVFNPPNGTHSWPYMNQOLVAMKADI 300

RESULT 15

US-08-997-080-36

Sequence 36, Application US/08997080

Patent No. 5968524

GENERAL INFORMATION:

APPLICANT: WATSON, JAMES D.

APPLICANT: TAN, PAUL L.O.

TITLE OF INVENTION: METHODS AND COMPOUNDS FOR THE TREATMENT OF IMMUNOLOGICALLY-

NUMBER OF SEQUENCES: 194

CORRESPONDENCE ADDRESS:

ADDRESSER: Law Offices of Ann W. Speckman

STREET: 2601 Elliott Avenue, Suite 4185

CITY: Seattle

STATE: WA

COUNTRY: USA

ZIP: 98121

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/997,080

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Sleath, Janet

REGISTRATION NUMBER: 37,007

REFERENCE/DOCKET NUMBER: 11000.1007

TELECOMMUNICATION INFORMATION:

TELEPHONE: 206-269-0565

TELEFAX: 206-269-0563

INFORMATION FOR SEQ ID NO: 36:

SEQUENCE CHARACTERISTICS:

LENGTH: 333 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-997-080-36

Query Match 33.8%; Score 548; DB 1; Length 333;

Best Local Similarity 39.2%; Pred. No. 1e-40;

Matches 118; Conservative 41; Mismatches 128; Indels 14; Gaps 5;

Db 6 ALIRALMIALSFLGSLGVAVAEPTAKAAPYENLMVPSPSMGRIPIVAPFLAGSHAYILL 65

Db 24 AVIGTALLAGLVGVDDTAIAVAFSKPGLPYEYLQVPSPSMGRIKIQVPOGGPHAYILL 83

Db 66 DAFNAGPDVSNWVTAGNANMTLAGKGISVYVAPAGAVSMYTMEDQ--DQSKQ-----WDT 118

Db 84 DGLRAQEDYNGMDINTPAFEERYHSGLSVIMPVGGQSSFTTDMYQPSQSNQNTTYKMET 143

Db 119 FLSAEI.PDMLAANRGLAPGHAIVAGAAQGGYGAMALAAFHEDRFAGSGNSGFLYPSNTT 178

Db 144 FLTQEMPSTWQANKNVLPGTNAIVAGLSWSGSSALIIASTYTPQOFPYAASLSGFLNBSG 203

Db 179 TNGAIAAGMOQFGGVDTNGMGAPQLGRWKMDPVTASLLAQNTRVYMW-----SPNPG 234

Db 204 WPTMIGLANWDSGGYANANSMWGPSTDPAAWKNDPMVOI PRLVANNTRIMWYCGNGAIPNL 263

Db 235 GASD-PAAMIGQTAAEMNSRMFTYQYRSVGHNHGFDPASGDNGSWAPQLGAMSDI 293

Db	264	GSDNIPAKL--ESLTLSTNEIFONTYAASGGRGVFNPPNGTSHWPYNQOLVANKPD	321
Qy	294	I 294	
Db	322	I 322	

Search completed: April 14, 2006, 17:37:57
Job time : 28.6477 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 14, 2006, 18:27:13 ; Search time 104.926 Seconds
(without alignments)
1190.652 Million cell updates/sec

Title: US-10-620-246-42

Perfect score: 1621
Sequence: 1 MKGRSALLRALMTALISFGL.....WGSMAPQLGMSGDIVGAIR 299

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Published Applications_AA_Main:
1: /cgn2_6/prodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/prodata/1/pubpaa/US08_PUBCOMB.pep:*
3: /cgn2_6/prodata/1/pubpaa/US09_PUBCOMB.pep:*
4: /cgn2_6/prodata/1/pubpaa/US10_PUBCOMB.pep:*
5: /cgn2_6/prodata/1/pubpaa/US10_PUBCOMB.pep:*
6: /cgn2_6/prodata/1/pubpaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1621	100.0	299	US-09-791-171-42	Sequence 42, Appl
2	1621	100.0	299	US-10-620-246-42	Sequence 42, Appl
3	1616	99.7	299	US-10-695-155-166	Sequence 166, App
4	1604.5	99.0	298	US-09-804-980-42	Sequence 42, Appl
5	1335	82.4	252	US-10-481-563A-14	Sequence 14, Appl
6	564.5	34.8	340	US-10-332-512A-3	Sequence 3, Appl
7	561.5	34.6	340	US-09-880-505-37	Sequence 37, Appl
8	561.5	34.6	340	US-10-051-643-37	Sequence 37, Appl
9	548	33.8	333	US-09-880-505-36	Sequence 36, Appl
10	548	33.8	333	US-10-051-643-36	Sequence 36, Appl
11	548	33.8	333	US-10-282-122A-64050	Sequence 64050, A
12	532	32.8	338	US-09-880-505-32	Sequence 32, Appl
13	532	32.8	338	US-09-880-505-34	Sequence 34, Appl
14	532	32.8	338	US-10-051-643-32	Sequence 32, Appl
15	532	32.8	338	US-10-051-643-34	Sequence 34, Appl
16	532	32.8	338	US-10-329-087-39	Sequence 39, Appl
17	532	32.8	338	US-10-203-562-4	Sequence 4, Appl
18	532	32.8	338	US-10-332-512A-1	Sequence 1, Appl
19	532	32.8	338	US-10-695-155-162	Sequence 162, App
20	523	32.3	330	US-09-880-505-30	Sequence 30, Appl
21	523	32.3	330	US-10-051-643-30	Sequence 30, Appl
22	522	32.2	333	US-10-329-087-37	Sequence 37, Appl
23	521	32.1	333	US-10-329-087-35	Sequence 35, Appl
24	519	32.0	325	US-09-805-427A-2	Sequence 2, Appl
25	519	32.0	325	US-09-880-505-33	Sequence 33, Appl
26	519	32.0	325	US-09-952-554-1	Sequence 1, Appl
27	519	32.0	325	US-09-872-505-2	Sequence 2, Appl

28	519	32.0	325	4	US-10-051-643-33	Sequence 33, Appl
29	519	32.0	325	4	US-10-689-921-14	Sequence 14, Appl
30	519	32.0	325	4	US-10-332-512A-2	Sequence 2, Appl
31	519	32.0	325	5	US-10-695-155-161	Sequence 161, App
32	518	32.0	326	3	US-09-880-505-43	Sequence 43, Appl
33	518	32.0	326	4	US-10-051-643-43	Sequence 43, Appl
34	518	32.0	326	4	US-10-205-979-29	Sequence 29, Appl
35	514.5	31.7	334	3	US-09-880-505-45	Sequence 45, Appl
36	514.5	31.7	334	4	US-10-051-643-45	Sequence 45, Appl
37	514.5	31.7	334	4	US-10-205-979-31	Sequence 31, Appl
38	513	31.6	295	4	US-10-332-512A-31	Sequence 31, Appl
39	513	31.6	328	4	US-10-329-087-43	Sequence 43, Appl
40	512.5	31.6	1016	4	US-10-369-983-18	Sequence 18, Appl
41	512	31.6	323	3	US-09-880-505-35	Sequence 35, Appl
42	512	31.6	323	4	US-10-051-643-35	Sequence 35, Appl
43	506	31.2	403	3	US-09-791-171-173	Sequence 173, App
44	506	31.2	403	3	US-09-805-427A-4	Sequence 4, Appl
45	506	31.2	403	3	US-09-804-980-173	Sequence 173, App

ALIGNMENTS

```
RESULT 1
US-09-791-171-42
; Sequence 42, Application US/09791171
; Patent No. US20020094336A1
; GENERAL INFORMATION:
; APPLICANT: ANDERSEN, Peter
; APPLICANT: NIELSEN, Rikke
; APPLICANT: OSTTINGER, Thomas
; APPLICANT: RASMUSSEN, Peter Birk
; APPLICANT: ROSENKRANDS, Ida
; APPLICANT: WEIDENKAR, Karin
; APPLICANT: FIORIO, Walter
; TITLE OF INVENTION: NUCLEIC ACIDS FRAGMENTS AND POLYPEPTIDE FRAGMENTS
; TITLE OF INVENTION: DERIVED FROM M. TUBERCULOSIS
; FILE REFERENCE: 670001-2002.1
; CURRENT APPLICATION NUMBER: US/09/791,171
; CURRENT FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 2001-02-20
; PRIOR FILING DATE: 1998-03-30
; PRIOR FILING DATE: 1998-03-30
; PRIOR APPLICATION NUMBER: 0376/97
; PRIOR FILING DATE: 1997-04-02
; PRIOR APPLICATION NUMBER: 1277/97
; PRIOR FILING DATE: 1997-11-10
; PRIOR APPLICATION NUMBER: 60/044,624
; PRIOR FILING DATE: 1997-04-18
; PRIOR APPLICATION NUMBER: 60/070,488
; PRIOR FILING DATE: 1998-01-05
; NUMBER OF SEQ ID NOS: 173
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 42
; LENGTH: 299
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-09-791-171-42

Query Match      100.0%; Score 1621; DB 3; Length 299;
Best Local Similarity 100.0%; Pred. No. 2.2e-130;
Matches 299; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY	1	MKGRSALLRALMTALISFGLGVAAVAAEPTAKAPYENLMVPSPMGORDIPVAFIAGGPH	60
DB	1	MKGRSALLRALMTALISFGLGVAAVAAEPTAKAPYENLMVPSPMGORDIPVAFIAGGPH	60
QY	61	AVYLLDAFNAAGPDVSNMVTAGNMMTLAAGKISVVAAPAGGAYSMYTNWBDGSKQMDTFL	120
DB	61	AVYLLDAFNAAGPDVSNMVTAGNMMTLAAGKISVVAAPAGGAYSMYTNWBDGSKQMDTFL	120
QY	121	SAELPMTLANRGLAAGGAAVGAAGGCGAMALAAFHPRRPFAGSMGSLYPSNTTNN	180
DB	121	SAELPMTLANRGLAAGGAAVGAAGGCGAMALAAFHPRRPFAGSMGSLYPSNTTNN	180

; APPLICANT: Anderson, Peter
 ; TITLE OF INVENTION: M. tuberculosis Antigens
 ; FILE REFERENCE: 670001-2002.4
 ; CURRENT APPLICATION NUMBER: US/09/804,980
 ; CURRENT FILING DATE: 2001-03-12
 ; NUMBER OF SEQ ID NOS: 257
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 42
 ; LENGTH: 298
 ; TYPE: PR1
 ; ORGANISM: Mycobacterium tuberculosis
 US-09-804-980-42

Query Match 99.0%; Score 1604.5; DB 3; Length 298;
 Best Local Similarity 99.7%; Pred. No. 5,76-129;
 Matches 298; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 MKGRSALLRLMTALISFGLGVAVAAPPAKAPYENLMVPSMGRIIPVAFLAGGPH 60
 DB 1 MKGRSALLRLMTALISFGLGVAVAAPPAKAPYENLMVPSMGRIIPVAFLAGGPH 59
 QY 61 AVYLLDAFNAAGPVSNMVTAGNANTLAGKISVVAAGAGYSMTTWEDGSKQMDTFL 120
 DB 60 AVYLLDAFNAAGPVSNMVTAGNANTLAGKISVVAAGAGYSMTTWEDGSKQMDTFL 119
 QY 121 SAELPDLTANRGLAPGHAAGVGAAGGAGMALAAHPDRFGAGMSGFLYPSNTTNN 180
 DB 120 SAELPDLTANRGLAPGHAAGVGAAGGAGMALAAHPDRFGAGMSGFLYPSNTTNN 179
 QY 181 GAIAAGQOFGVDYTNMGWAPOLGRMKMDPVVHASLLAQNNTTRVWVMSPTNPGASDPA 240
 DB 180 GAIAAGQOFGVDYTNMGWAPOLGRMKMDPVVHASLLAQNNTTRVWVMSPTNPGASDPA 239
 QY 241 AMGTGTAAAGNSRMFTNQRYSVGHGHPDPASGNGMSGVAPOLGAMSGDIVGAIR 299
 DB 240 AMGTGTAAAGNSRMFTNQRYSVGHGHPDPASGNGMSGVAPOLGAMSGDIVGAIR 298

RESULT 5

US-10-481-563A-14
 ; Sequence 14, Application US/10481563A
 ; Publication No. US20050084904A1
 ; GENERAL INFORMATION:
 ; APPLICANT: NEW YORK UNIVERSITY
 ; TITLE OF INVENTION: MYCOBACTERIAL PROTEINS AS EARLY ANTIGENS FOR SERODIAGNOSIS AND V
 ; FILE REFERENCE: 32004-173355
 ; CURRENT APPLICATION NUMBER: US/10/481,563A
 ; CURRENT FILING DATE: 2003-12-22
 ; PRIOR APPLICATION NUMBER: US 60/212,048
 ; PRIOR FILING DATE: 2000-06-16
 ; NUMBER OF SEQ ID NOS: 16
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 14
 ; LENGTH: 252
 ; TYPE: PR1
 ; ORGANISM: Mycobacterium tuberculosis
 US-10-481-563A-14

Query Match 82.4%; Score 1335; DB 5; Length 252;
 Best Local Similarity 94.7%; Pred. No. 5,4e-106;
 Matches 252; Conservative 0; Mismatches 0; Indels 14; Gaps 5;

QY 34 APYENLMVPSMGRIIPVAFLAGGPHAVYLLDAFNAAGPVSNMVTAGNANTLAGKIS 93
 DB 1 APYENLMVPSMGRIIPVAFLAGGPHAVYLLDAFNAAGPVSNMVTAGNANTLAGKIS 59
 QY 94 VVAAGAGYSMTTWEDGSKQMDTFLSAELPDLTANRGLAPGHAAGVGAAGGAGMALAA 153
 DB 60 VVAAGAGYSMTTWEDGSKQMDTFLSAELPDLTANRGLAPGHAAGVGAAGGAGMALAA 109
 QY 154 LAAPHDRFGAGMSGFLYPSNTTNGAIAAGQOFGVDYTNMGWAPOLGRMKMDPV 213
 DB 110 -AAFHDRFGAGMSGFLYPSNTTNGAIAAGQOFGVDYTNMGWAPOLGRMKMDPV 168

QY 214 VHASLLAQNNTTRVWVMSPTNPGASDPAAMIQTAAAGNSRMFTNQRYSVGHGHPDP 273
 DB 169 -HASLLAQNNTTRVWVMSPTNPGASDPAAMIQTAAAGNSRMFTNQRYSVGHGHPDP 227
 QY 274 ASGDNWGSWAPOLGAMSGDIVGAIR 299
 DB 228 -SGDNWGSWAPOLGAMSGDIVGAIR 252

RESULT 6

US-10-332-512A-3
 ; Sequence 3, Application US/10332512A
 ; Publication No. US20040180056A1
 ; GENERAL INFORMATION:
 ; APPLICANT: ORME, Ian M.
 ; APPLICANT: BRISLE, John T.
 ; TITLE OF INVENTION: MID-LIFE VACCINE AND METHODS FOR BOOSTING ANTI-MYCOBACTERIAL IMM
 ; FILE REFERENCE: 38861-166292
 ; CURRENT APPLICATION NUMBER: US/10/332,512A
 ; CURRENT FILING DATE: 2003-01-10
 ; PRIOR APPLICATION NUMBER: PCT/US01/21717
 ; PRIOR FILING DATE: 2001-07-10
 ; PRIOR APPLICATION NUMBER: US 60/217,646
 ; PRIOR FILING DATE: 2000-07-10
 ; NUMBER OF SEQ ID NOS: 31
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 3
 ; LENGTH: 340
 ; TYPE: PR1
 ; ORGANISM: Mycobacterium tuberculosis
 US-10-332-512A-3

Query Match 34.8%; Score 564.5; DB 4; Length 340;
 Best Local Similarity 41.6%; Pred. No. 6,3e-40;
 Matches 131; Conservative 38; Mismatches 119; Indels 27; Gaps 8;

QY 4 RSA---LRLMTIA---LSFGL---GGVAVAEPTAKAPYENLMVPSMGRIIPV 52
 DB 11 RSNATTLRLMTIAAGAVLYVGLVGFPGPATAGAFSRGLVEYLVQVPSASGRDIKV 70
 QY 53 AFLAGPHAVYLLDAFNAAGPVSNMVTAGNANTLAGKISVVAAGAGYSMTTWEDG 112
 DB 71 OFGGGPHAVYLLDGAFAODDYNMGMDINTPAFEREYYSGLSVIMPVGGSSFTYDWQPS 130
 QY 113 SK-----QMDTFLSAELPDLTANRGLAPGHAAGVGAAGGAGMALAAFPDRGFA 165
 DB 131 QSNQNTYTKMETFLTRFEMPAMLOANKGVSPGNAAGLSMSGASALITLAAYTPOQFPA 190
 QY 166 GMSGFLYPSNTTNGAIAAGQOFGVDYTNMGWAPOLGRMKMDPVVHASLLAQNNT 225
 DB 191 ASLSGFLNBSGFWPLILGLAMNDSSGYANNSMGSSDPAWGRNDPMVQIPRLVANNT 250
 QY 226 VVWV---SPTNPGASD-PAAMI-GQTAAAGNSRMFTNQRYSVGHGHPDPASGDN 279
 DB 251 IVVYCGNGFPSDLDGDNIPAKFLEGLT---LRTNQTFRDYADAGGRNVGFNPPNGTHS 307
 QY 280 WGSWAPOLGAMSGDI 294
 DB 308 WPYWNEQLVAMKADI 322

RESULT 7

US-09-880-505-37
 ; Sequence 37, Application US/09880505
 ; Publication No. US20030007976A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Watson, James D.
 ; APPLICANT: Tan, Paul L.J.
 ; APPLICANT: Prestidge, Ross
 ; TITLE OF INVENTION: Methods and Compounds for the Treatment
 ; FILE REFERENCE: 11000.1007c2

```

CURRENT APPLICATION NUMBER: US/09/880,505
CURRENT FILING DATE: 2001-06-13
PRIOR APPLICATION NUMBER: US 09/324,542
PRIOR FILING DATE: 1999-06-02
PRIOR APPLICATION NUMBER: US 08/997,080
PRIOR FILING DATE: 1997-12-23
NUMBER OF SEQ ID NOS: 194
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 37
LENGTH: 340
TYPE: PRT
ORGANISM: Mycobacterium tuberculosis
US-09-880-505-37

```

[illegible]

```

RESULT 8
US-10-051-643-37
; Sequence 37, Application US/10051643
; Publication No. US20020197265A1
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Tan, Paul L. J.
; TITLE OF INVENTION: Methods and Compounds for the Treatment
; TITLE OF INVENTION: of Immunologically-Mediated Diseases of the Respiratory
; TITLE OF INVENTION: System using Mycobacterium Vaccae
; FILE REFERENCE: 11000.1008C2
; CURRENT FILING DATE: US/10/051,643
; PRIOR FILING DATE: 2002-01-18
; PRIOR APPLICATION NUMBER: US09/156,181
; PRIOR FILING DATE: 1998-09-17
; PRIOR APPLICATION NUMBER: US 08/996,624
; PRIOR FILING DATE: 1997-12-23
; NUMBER OF SEQ ID NOS: 208
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 37
; LENGTH: 340
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-10-051-643-37

```

```

Query Match          34.6%; Score 561.5; DB 4; Length 340;
Best Local Similarity 41.3%; Pred. NO. 1.2e-39;
Matches 130; Conservative 39; Mismatches 119; Indels 27; Gaps 8;

4 RSA---LIRALMIAA-----LSFGL-----GGVAVVAAPFAAKAAPYINLNVPSBSMDRIDPV 52

```

Db 11 R\$AAITLRRRAIAAMGAVLVYGLVGTGPGPATAGAF\$R\$GLPVELVQ\$P\$ASMG\$RDIK\$ 70

Qy 53 AFLAGGPAVYLLDAFNAQPDVSNVVTGNAMNTLAGKISVYAPGAGAYSMTYNQDS 112

Db 71 QFGGGPFAVYLLDGLRAODDDYNGMDINTPAREEYYSQ\$GL\$VIM\$P\$GQSS\$P\$TIDWYQ\$S 130

Qy 113 SK-----QMDPL\$A\$EL\$PMDL\$ANRGLA\$P\$G\$H\$A\$V\$G\$A\$V\$G\$Y\$G\$M\$A\$L\$A\$F\$P\$D\$R\$F\$G\$A 165

Db 131 Q\$NG\$N\$Q\$Y\$Y\$K\$E\$T\$E\$F\$T\$R\$E\$M\$P\$A\$W\$L\$Q\$N\$K\$G\$V\$S\$P\$T\$G\$N\$A\$V\$G\$L\$S\$M\$G\$G\$S\$A\$L\$I\$A\$Y\$T\$P\$Q\$P\$P\$Y\$A 190

Qy 166 G\$M\$S\$G\$F\$L\$P\$S\$N\$T\$T\$G\$A\$I\$A\$A\$G\$Q\$F\$G\$Y\$D\$T\$N\$G\$M\$G\$A\$P\$Q\$L\$R\$K\$M\$H\$P\$W\$Y\$A\$S\$L\$A\$O\$N\$T\$R 225

Db 191 A\$S\$L\$G\$F\$L\$P\$B\$E\$C\$R\$W\$T\$T\$L\$G\$L\$A\$N\$D\$G\$Y\$A\$N\$S\$M\$G\$S\$S\$D\$P\$A\$K\$R\$N\$D\$P\$W\$Q\$I\$P\$L\$V\$A\$N\$N\$R 250

Qy 226 V\$W\$V\$-----S\$P\$T\$N\$P\$A\$S\$D\$-P\$A\$M\$I-\$Q\$Y\$A\$E\$A\$M\$N\$S\$R\$E\$F\$T\$N\$Q\$Y\$S\$V\$G\$H\$N\$G\$H\$F\$D\$P\$A\$S\$G\$D\$N\$G 279

Db 251 I\$W\$Y\$G\$N\$G\$T\$P\$S\$D\$L\$G\$D\$N\$I\$P\$A\$K\$L\$E\$G\$L\$T\$--L\$R\$T\$N\$O\$T\$F\$R\$D\$Y\$A\$A\$D\$G\$R\$N\$G\$V\$F\$N\$P\$P\$N\$G\$T\$H\$S 307

Qy 280 W\$G\$N\$A\$P\$Q\$L\$G\$A\$M\$S\$G\$D\$I 294

Db 308 W\$P\$Y\$N\$E\$Q\$L\$V\$A\$M\$K\$A\$D\$I 322

```

RESULT 9
US-09-880-505-36
; Sequence 36, Application US/09880505
; Publication No. US20030007976A1
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Tan, Paul L.J.
; APPLICANT: Prestidge, Ross
; TITLE OF INVENTION: Methods and Compounds for the Treatment
; TITLE OF INVENTION: of Immunologically-Mediated Skin Disorders
; FILE REFERENCE: 11000.1007C2
; CURRENT APPLICATION NUMBER: US/09/880,505
; CURRENT FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 09/324,542
; PRIOR FILING DATE: 1999-06-02
; PRIOR APPLICATION NUMBER: US 08/997,080
; PRIOR FILING DATE: 1997-12-23
; NUMBER OF SEQ ID NOS: 194
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 36
; LENGTH: 333
; TYPE: PRT
; ORGANISM: Mycobacterium leprae
US-09-880-505-36

```

Query Match	33.8%	Score 548	DB 3	Length 333
Beet Local Similarity	39.2%	Pred. No. 1.7e-38		
Matches	118	Conservative 41	Mismatches 128	Indels 14
			Gaps	5
QY	6	ALRLALMTAAASFGGAGVAAPEPAKAPENYLNWPPSSMGDRDIPVAFLAGGPHAVYLL	65	
DB	24	AVTGTAIIAGLVGVGDTRIAVAFSKPGELPVETLYOVPSMSMGHDITIQGGGQHAYLL	83	
QY	66	DAFNAGPVSNNVVTGNNAMNTLAGKGISVVAAGAGYSMTYTWEQ--DGSKQ-----WDT	118	
DB	84	DGLRAQEDYNGMDINTPRAEERYTHSGLSYIMVVGQSSFSYNNWYQPSQNGCHYTKMET	143	
QY	119	FLSAELPMLAANRGLAPGHAAGAAGGAGYAMALAAHPDRFCFAGSMGFLYPSNTT	178	
DB	144	FLTGEMPSMLQNNKNVLPFGNAAVGLSMGSSALILASYYPOQFPYAASLGFLNPSEGW	203	
QY	179	TNGALIAAGQQQFGVDYTGAMGAPQLGRKKMHDPMVHASLLAQNNTRVVW----	234	
DB	204	WPTMIGLMMNDSGGYNAISMWGPSTDPAPKRRDPMWQIPRLVAANTRIVYCGNGAPNEL	263	
QY	225	GASD-PAAMIGTAAAMGSMKRFYQYASVGGHNHPPDPASGDNCGSMAPOLGMSGD	253	
DB	264	GGDNIPAEFL--ESLTJLTSTNELFQYTYAASGGKNGFNFPNGTSHWPYNNQOLVAMKPD	321	

QY 294 I 294
DB 322 I 322

RESULT 10
US-10-051-643-36
; Sequence 36, Application US/10051643
; Publication No. US20020197265A1
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Tan, Paul L. J.
; TITLE OF INVENTION: Methods and Compounds for the Treatment
; TITLE OF INVENTION: of Immunologically-Mediated Diseases of the Respiratory
; FILE REFERENCE: 11000.1008c2
; CURRENT APPLICATION NUMBER: US/10/051,643
; PRIOR FILING DATE: 2002-01-18
; PRIOR APPLICATION NUMBER: US09/156,181
; PRIOR FILING DATE: 1998-09-17
; PRIOR APPLICATION NUMBER: US 08/996,624
; PRIOR FILING DATE: 1997-12-23
; NUMBER OF SEQ ID NOS: 208
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 36
; LENGTH: 333
; TYPE: PR1
; ORGANISM: Mycobacterium leprae
US-10-051-643-36

Query Match 33.8%; Score 548; DB 4; Length 333;
Best Local Similarity 39.2%; Pred. No. 1,7e-38;
Matches 118; Conservative 41; Mismatches 128; Indels 14; Gaps 5;

QY 6 ALBALMIALSLGCGVAVAAEPTAKAAYENLAMPSPSGRIPVAFLAGPHAYLL 65
DB 24 AVIGTALLAGLVGVGTALAVAFSKPGLPVEYIQVPSPSMGHDIKIQFGGGOHAYLL 83
QY 66 DATNAGPDVSNWYTAGNANTLAGKISVYAPAGAYSMYTNMQ--DGSKO-----NDT 118
DB 84 DGLRAQEDYNGMDINTPAFEERYHSGLSVIMPVGGSSFTSNMYQPSQNGQHTTYMET 143
QY 119 FLASLEPDMLANRGLAPGHAAGAAGGAGMALAFAHDFRFGAGSMGFLYPENTT 178
DB 144 FLTQEWPSWLOAKNVLPTGNAAVGLSMGSSALILASYYPOQFPYAAISGFLNPBEGW 203
QY 179 TNGAIAAGQOQFGVDYTNMGWAPQLGRWTHDPWHASLLAONNTVWYV---SPTNP 234
DB 204 WPTMIGLANNDSCGYANANSMGSPSTDPAWKRNDPMVQIPRLVANNTIRIWWYCGGAGNEL 263
QY 235 GASD-PAAMIGQTAEMANGSRMFPYNOYRSVGHNGHDFPASGDNMGWSAPOLGAMSGD 293
DB 264 GSDNIPAKFL--ESLTLSTNEIFONTYASGGRNGVFNFPNGTHSPYNNQOLVAKPD 321
QY 294 I 294
DB 322 I 322

RESULT 11
US-10-282-122A-64050
; Sequence 64050, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zykkind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant

APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA.034A
CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: Patentin version 3.1
SEQ ID NO 64050
LENGTH: 333
TYPE: PR1
ORGANISM: Mycobacterium leprae
US-10-282-122A-64050

Query Match 33.8%; Score 548; DB 4; Length 333;
Best Local Similarity 39.2%; Pred. No. 1,7e-38;
Matches 118; Conservative 41; Mismatches 128; Indels 14; Gaps 5;

QY 6 ALBALMIALSLGCGVAVAAEPTAKAAYENLAMPSPSGRIPVAFLAGPHAYLL 65
DB 24 AVIGTALLAGLVGVGTALAVAFSKPGLPVEYIQVPSPSMGHDIKIQFGGGOHAYLL 83
QY 66 DATNAGPDVSNWYTAGNANTLAGKISVYAPAGAYSMYTNMQ--DGSKO-----NDT 118
DB 84 DGLRAQEDYNGMDINTPAFEERYHSGLSVIMPVGGSSFTSNMYQPSQNGQHTTYMET 143
QY 119 FLASLEPDMLANRGLAPGHAAGAAGGAGMALAFAHDFRFGAGSMGFLYPENTT 178
DB 144 FLTQEWPSWLOAKNVLPTGNAAVGLSMGSSALILASYYPOQFPYAAISGFLNPBEGW 203
QY 179 TNGAIAAGQOQFGVDYTNMGWAPQLGRWTHDPWHASLLAONNTVWYV---SPTNP 234
DB 204 WPTMIGLANNDSCGYANANSMGSPSTDPAWKRNDPMVQIPRLVANNTIRIWWYCGGAGNEL 263
QY 235 GASD-PAAMIGQTAEMANGSRMFPYNOYRSVGHNGHDFPASGDNMGWSAPOLGAMSGD 293
DB 264 GSDNIPAKFL--ESLTLSTNEIFONTYASGGRNGVFNFPNGTHSPYNNQOLVAKPD 321
QY 294 I 294
DB 322 I 322

RESULT 12
US-09-880-505-32
; Sequence 32, Application US/09880505
; Publication No. US20030007976A1
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Tan, Paul L. J.
; APPLICANT: Prestidge, Ross

```
/ TITLE OF INVENTION: Methods and Compounds for the Treatment
/ TITLE OF INVENTION: of Immunologically-Mediated Skin Disorders
/ FILE REFERENCE: 11000.1007C2
/ CURRENT APPLICATION NUMBER: US/09/880,505
/ CURRENT FILING DATE: 2001-06-13/324,542
/ PRIOR APPLICATION NUMBER: US 09/324,542
/ PRIOR FILING DATE: 1999-06-02
/ PRIOR APPLICATION NUMBER: US 08/997,080
/ PRIOR FILING DATE: 1997-12-23
/ NUMBER OF SEQ ID NOS: 194
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 32
/ LENGTH: 338
/ TYPE: PRF
/ ORGANISM: Mycobacterium tuberculosis
US-09-880-505-32
```

```
Query Match 32.8%; Score 532; DB 3; Length 338;
Best Local Similarity 40.1%; Pred. No. 4.1e-37;
Matches 121; Conservative 37; Mismatches 126; Indels 18; Gaps 5;
```

```
QY 14 AALSFGT-----GGVAVAABPTAKAAYENILMVPSPSGMDIPVAFLAGSPH--AVYLLDA 67
DB 25 AALVSLVGAAGVGTATAGAFSRPGLPVEYLQVPSPSMGDKIQOFSGGANSPLATLDDG 84
QY 68 FNAGPDVSNWVTAGNANMTLAGKISVAPAGAYMYTNMBODG-----SKQWDITPL 120
DB 85 LRAQDDPSGMDINTPAFEWYDQSGLSVMPVGQSSFYSDWYOPACGKACQYTKMETFL 144
QY 121 SAELPMTLANRGLACGHAAGVAAAGCGYGAALAAFHDPFGPAGSMGFLYPSNTTNN 180
DB 145 TSELPGMLQANRRHVKFTGSAVVGLSMAASSALTTLATYHPOQFYVAGAMSGLLDPSQAMGP 204
QY 181 GAIAAGMOQFGVDITNGMNGAPQLGRKMHDPVNHASLLAQNNTYVWVW-----SPTNPGA 236
DB 205 TLIGLMDAGAGYKASDMPKEDPAWQKNDPLNNGKLIANNTRYVWYCGNGKPSDLGG 264
QY 237 SDPAMIGQTAEAMGNSRMFYNQYRSVGGHGHFDPFASGDNMGSWAPQLGAMSGDIYV 296
DB 265 NNLPATFLEGFVYRTSNIK-FQDAVYNAGGCHNGVDFPDSGTHSWETWGAQLNAMKEDLQR 323
QY 297 AI 298
DB 324 AL 325
```

```
RESULT 13
US-09-880-505-34
/ Sequence 34, Application US/09880505
/ Publication No. US20030007976A1
/ GENERAL INFORMATION:
/ APPLICANT: Watson, James D.
/ APPLICANT: Tan, Paul L.J.
/ APPLICANT: Prestidge, Ross
/ TITLE OF INVENTION: Methods and Compounds for the Treatment
/ TITLE OF INVENTION: of Immunologically-Mediated Skin Disorders
/ FILE REFERENCE: 11000.1007C2
/ CURRENT APPLICATION NUMBER: US/09/880,505
/ CURRENT FILING DATE: 2001-06-13
/ PRIOR APPLICATION NUMBER: US 09/324,542
/ PRIOR FILING DATE: 1999-06-02
/ PRIOR APPLICATION NUMBER: US 08/997,080
/ PRIOR FILING DATE: 1997-12-23
/ NUMBER OF SEQ ID NOS: 194
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 34
/ LENGTH: 338
/ TYPE: PRF
/ ORGANISM: Mycobacterium bovis
US-09-880-505-34
```

```
Query Match 32.8%; Score 532; DB 3; Length 338;
Best Local Similarity 40.1%; Pred. No. 4.1e-37;
```

```
Matches 121; Conservative 37; Mismatches 126; Indels 18; Gaps 5;
QY 14 AALSFGT-----GGVAVAABPTAKAAYENILMVPSPSGMDIPVAFLAGSPH--AVYLLDA 67
DB 25 AALVSLVGAAGVGTATAGAFSRPGLPVEYLQVPSPSMGDKIQOFSGGANSPLATLDDG 84
QY 68 FNAGPDVSNWVTAGNANMTLAGKISVAPAGAYMYTNMBODG-----SKQWDITPL 120
DB 85 LRAQDDPSGMDINTPAFEWYDQSGLSVMPVGQSSFYSDWYOPACGKACQYTKMETFL 144
QY 121 SAELPMTLANRGLACGHAAGVAAAGCGYGAALAAFHDPFGPAGSMGFLYPSNTTNN 180
DB 145 TSELPGMLQANRRHVKFTGSAVVGLSMAASSALTTLATYHPOQFYVAGAMSGLLDPSQAMGP 204
QY 181 GAIAAGMOQFGVDITNGMNGAPQLGRKMHDPVNHASLLAQNNTYVWVW-----SPTNPGA 236
DB 205 TLIGLMDAGAGYKASDMPKEDPAWQKNDPLNNGKLIANNTRYVWYCGNGKPSDLGG 264
QY 237 SDPAMIGQTAEAMGNSRMFYNQYRSVGGHGHFDPFASGDNMGSWAPQLGAMSGDIYV 296
DB 265 NNLPATFLEGFVYRTSNIK-FQDAVYNAGGCHNGVDFPDSGTHSWETWGAQLNAMKEDLQR 323
QY 297 AI 298
DB 324 AL 325
```

```
RESULT 14
US-10-051-643-32
/ Sequence 32, Application US/10051643
/ Publication No. US20020197265A1
/ GENERAL INFORMATION:
/ APPLICANT: Watson, James D.
/ APPLICANT: Tan, Paul L. J.
/ TITLE OF INVENTION: Methods and Compounds for the Treatment
/ TITLE OF INVENTION: of Immunologically-Mediated Diseases of the Respiratory
/ FILE REFERENCE: 11000.1008C2
/ CURRENT APPLICATION NUMBER: US/10/051,643
/ CURRENT FILING DATE: 2002-01-18
/ PRIOR APPLICATION NUMBER: US09/156,181
/ PRIOR FILING DATE: 1998-09-17
/ PRIOR APPLICATION NUMBER: US 08/996,624
/ PRIOR FILING DATE: 1997-12-23
/ NUMBER OF SEQ ID NOS: 208
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 32
/ LENGTH: 338
/ TYPE: PRF
/ ORGANISM: Mycobacterium tuberculosis
US-10-051-643-32
```

```
Query Match 32.8%; Score 532; DB 4; Length 338;
Best Local Similarity 40.1%; Pred. No. 4.1e-37;
Matches 121; Conservative 37; Mismatches 126; Indels 18; Gaps 5;
QY 14 AALSFGT-----GGVAVAABPTAKAAYENILMVPSPSGMDIPVAFLAGSPH--AVYLLDA 67
DB 25 AALVSLVGAAGVGTATAGAFSRPGLPVEYLQVPSPSMGDKIQOFSGGANSPLATLDDG 84
QY 68 FNAGPDVSNWVTAGNANMTLAGKISVAPAGAYMYTNMBODG-----SKQWDITPL 120
DB 85 LRAQDDPSGMDINTPAFEWYDQSGLSVMPVGQSSFYSDWYOPACGKACQYTKMETFL 144
QY 121 SAELPMTLANRGLACGHAAGVAAAGCGYGAALAAFHDPFGPAGSMGFLYPSNTTNN 180
DB 145 TSELPGMLQANRRHVKFTGSAVVGLSMAASSALTTLATYHPOQFYVAGAMSGLLDPSQAMGP 204
QY 181 GAIAAGMOQFGVDITNGMNGAPQLGRKMHDPVNHASLLAQNNTYVWVW-----SPTNPGA 236
DB 205 TLIGLMDAGAGYKASDMPKEDPAWQKNDPLNNGKLIANNTRYVWYCGNGKPSDLGG 264
QY 237 SDPAMIGQTAEAMGNSRMFYNQYRSVGGHGHFDPFASGDNMGSWAPQLGAMSGDIYV 296
```

Db 265 NNLPAKFLBGFVTRTSNIK-FODAYNAGGHNVPDPDSGTHSWYGAQJLNAKPDQR 323
QY 297 AI 298
Db 324 AL 325

RESULT 15

US-10-051-643-34
; Sequence 34, Application US/10051643
; Publication No. US20020197265A1
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Tan, Paul L. J.
; TITLE OF INVENTION: Methods and Compounds for the Treatment
; TITLE OF INVENTION: of Immunologically-Mediated Diseases of the Respiratory
; TITLE OF INVENTION: System using Mycobacterium Vaccine
; FILE REFERENCE: 11000.1008C2
; CURRENT APPLICATION NUMBER: US/10/051,643
; CURRENT FILING DATE: 2002-01-18
; PRIOR APPLICATION NUMBER: US09/156,181
; PRIOR FILING DATE: 1998-09-17
; PRIOR APPLICATION NUMBER: US 08/996,624
; PRIOR FILING DATE: 1997-12-23
; NUMBER OF SEQ ID NOS: 208
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 34
; LENGTH: 338
; TYPE: PRT
; ORGANISM: Mycobacterium bovis
US-10-051-643-34

Query Match 32.8%; Score 532; DB 4; Length 338;
Best Local Similarity 40.1%; Pred. No. 4.1e-37;
Matches 121; Conservative 37; Mismatches 126; Indels 18; Gaps 5;

QY 14 AALSFGT---GGVVAABPTAKAPYENIMVPSMGRDIPVAFLAGPH--AVYLLDA 67
DB 25 AALVSGIVAGVGTATAGAFSRPGLPYEYLQVSPSMGRDIKVQFSGANSPALYLLDG 84
QY 68 FNAQPDVSNVVTAGNANTLAGKISVVAAPAGAYSMYTNWBDG-----SKQWDTFL 120
DB 85 LRAQDDPSGWDINTPAPEMTDOGLSYMPVGGSSFYSDWYOPACGKACQTYKMETFL 144
QY 121 SAELPMTLANRGLAPCGHAAYGAAGCYGAMALAAHPDRFGAGSMGSLYPSNTTN 180
DB 145 TSELPGWLQANRHVKPTGSAVVGISMASASLTALAIYHPOQFYVYAGAMSGLLDPSQAMP 204
QY 181 GATAAGNQCRGVDYNTGMCAPOLGRYKWHDPVWVASLDAQNNTRVWVW-----SPTNPGA 236
DB 205 TLIGLAMGDAAGYKASIDMKGPKEPAMQRDPLIANVGKLIANNTRVWYCGNGKPSDLGG 264
QY 237 SDPAMIGQTAEMNGSRMYNOYRSVGNHGHDPDPASGDNGCWAPOLGANSGLDIVG 296
DB 265 NNLPAKFLBGFVTRTSNIK-FODAYNAGGHNVPDPDSGTHSWYGAQJLNAKPDQR 323
QY 297 AI 298
DB 324 AL 325

Search completed: April 14, 2006, 18:40:37
Job time : 106.026 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using SW model

Run on: April 14, 2006, 18:37:38 ; Search time 12.8124 Seconds
(without alignments)
991.565 Million cell updates/sec

Title: US-10-620-246-42

Perfect score: 1621
Sequence: 1 MKRSMALRLMTALSLFGU.....WGSWAPQLGAMSGDIVGAIR 299

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 217505 seqs, 42489236 residues

Total number of hits satisfying chosen parameters: 217505

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_AA_New:*
1: /SID55/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
2: /SID55/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
3: /SID55/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
4: /SID55/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
5: /SID55/ptodata/2/pubpaa/US05_NEW_PUB.pep:*
6: /SID55/ptodata/2/pubpaa/US11_NEW_PUB.pep:*
7: /SID55/ptodata/2/pubpaa/US11_NEW_PUB.pep:*
8: /SID55/ptodata/2/pubpaa/US60_NEW_PUB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	176.5	10.9	331	US-11-079-463-9174	Sequence 9174, App
2	159	9.8	165	US-11-082-389-68	Sequence 68, App1
3	121.5	7.5	783	US-11-052-554A-157	Sequence 157, App
4	108.5	6.7	388	US-11-188-298-11967	Sequence 11967, A
5	107.5	6.6	490	US-11-188-298-19912	Sequence 19912, A
6	107.5	6.6	606	US-11-052-554A-163	Sequence 163, App
7	107.5	6.6	1053	US-11-052-554A-151	Sequence 151, App
8	106.5	6.6	1492	US-11-188-298-11858	Sequence 11858, A
9	106.5	6.6	1643	US-11-052-554A-172	Sequence 172, App
10	106	6.5	543	US-11-052-554A-162	Sequence 162, App
11	105.5	6.5	292	US-11-172-740-542	Sequence 542, App
12	105.5	6.5	486	US-11-188-298-17488	Sequence 17488, A
13	105.5	6.5	493	US-11-188-298-8448	Sequence 8448, App
14	105	6.5	618	US-11-052-554A-150	Sequence 150, App
15	103	6.4	319	US-11-087-099-4262	Sequence 4262, App
16	100.5	6.2	325	US-11-188-298-13202	Sequence 13202, A
17	100.5	6.2	801	US-11-052-554A-166	Sequence 166, App
18	100	6.2	914	US-11-052-554A-160	Sequence 160, App
19	100	6.2	3300	US-11-052-554A-133	Sequence 133, App
20	99	6.1	453	US-11-087-099-12200	Sequence 12200, A
21	98.5	6.1	342	US-11-087-099-4274	Sequence 4274, App
22	98.5	6.1	767	US-11-052-554A-154	Sequence 154, App
23	98	6.0	449	US-11-169-041-138	Sequence 138, App
24	98	6.0	564	US-10-517-939-318	Sequence 318, App
25	98	6.0	778	US-11-052-554A-144	Sequence 144, App

26	98	6.0	1306	US-11-052-554A-139	Sequence 139, App
27	97.5	6.0	1436	US-11-052-554A-140	Sequence 140, App
28	97	6.0	837	US-11-052-554A-159	Sequence 159, App
29	96.5	6.0	330	US-11-188-298-2825	Sequence 2825, App
30	96.5	6.0	330	US-11-188-298-21964	Sequence 21964, A
31	96.5	6.0	3157	US-11-052-554A-142	Sequence 142, App
32	95.5	5.9	273	US-11-172-740-539	Sequence 539, App
33	95	5.9	433	US-11-188-298-8416	Sequence 8416, App
34	95	5.9	615	US-11-052-554A-152	Sequence 152, App
35	95	5.9	3716	US-11-052-554A-141	Sequence 141, App
36	94.5	5.8	639	US-11-052-554A-165	Sequence 165, App
37	94.5	5.8	1538	US-11-052-554A-146	Sequence 146, App
38	94	5.8	355	US-11-222-451-2	Sequence 2, App1
39	93.5	5.8	373	US-11-096-568A-12718	Sequence 12718, A
40	93.5	5.8	416	US-11-096-568A-12717	Sequence 12717, A
41	93.5	5.8	1901	US-11-052-554A-135	Sequence 135, App
42	93	5.7	388	US-11-130-821-1	Sequence 1, App1
43	93	5.7	518	US-11-188-298-9414	Sequence 9414, App
44	93	5.7	749	US-11-052-554A-148	Sequence 148, App
45	92.5	5.7	480	US-11-188-298-14874	Sequence 14874, A

ALIGNMENTS

```

RESULT 1
US-11-079-463-9174
; Sequence 9174, Application US/11079463
; Publication No. US20060073161A1
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO BACTERIOIDES FR
; FILE REFERENCE: PATHO0-03DIV2
; CURRENT APPLICATION NUMBER: US/11/079,463
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/128,705
; PRIOR FILING DATE: 1999-04-09
; PRIOR APPLICATION NUMBER: US 09/540,209
; PRIOR FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 10444
; SEQ ID NO 9174
; LENGTH: 331
; TYPE: PRT
; ORGANISM: B. fragilis
US-11-079-463-9174

Query Match      10.9%; Score 176.5; DB 7; Length 331;
Best Local Similarity 30.2%; Pred. No. 2.0e-07;
Matches 64; Conservative 35; Mismatches 74; Indels 39; Gaps 10;

QY      1 MKRSMALRLMTALSLFGIGVAVAAEPYAKAPYENLWVPSPKGRDIPVAF---LA 56
        60 MKRNNFLACLTLFVPL-----SARVDTLWVKSNNKEVQVLTVPDVA 106
        57 GGPAA-----VYLIDANAGPDVSNWYTAGNANMTLAG-KGISVAPAGAVSYMTWEO 110
        107 LGRNNAACFVLYLHG--GGAHTWQIRPNLPEIDDKGILFVCC-DKDSHYMSPK 163
QY      111 DSKRWDTFLSABLPMTLAN-RGLAP-GGAAVGAAGGYGAMALAAFPDRFGAGSM 168
        164 NPARYETFSSELYVYIDRNYKTIADRKRAITGLSGMGHAMWLGIRHKDVGGAAGST 223
QY      169 SGFL-----YPSNTTGCATAAGQOQGVGVTVN 196
        224 SGGVDIRPFPKNWSMN-----KQJGELASN 248

RESULT 2
US-11-082-389-68
; Sequence 68, Application US/11082389
; Publication No. US20050244935A1
; GENERAL INFORMATION:

```

```

; APPLICANT: Pompejus, Markus
; APPLICANT: Krogset, Burthard
; APPLICANT: Schroder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Heberhauser, Gregor
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS
; TITLE OF INVENTION: INVOLVED IN MEMBRANE SYNTHESIS AND MEMBRANE
; FILE REFERENCE: BGI-13ICPCN
; CURRENT APPLICATION NUMBER: US/11/082,389
; PRIOR FILING DATE: 2005-03-16
; PRIOR APPLICATION NUMBER: US 09/603024
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 60/141031
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: US 60/143262
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: US 60/151281
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: DE 19930487.4
; PRIOR FILING DATE: 1999-07-01
; PRIOR APPLICATION NUMBER: DE 19930489.0
; PRIOR FILING DATE: 1999-07-01
; PRIOR APPLICATION NUMBER: DE 19931549.3
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931550.7
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19932134.5
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19941379.7
; PRIOR FILING DATE: 1999-08-31
; Remaining Prior Application data removed - See file Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 446
; SEQ ID NO 68
; LENGTH: 165
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
; US-11-082-389-68

Query Match          9.8%; Score 159; DB 7; Length 165;
Best Local Similarity 32.4%; Pred. No. 3,5e-06;
Matches 55; Conservative 25; Mismatches 70; Indels 20; Gaps 7;

QY      142 VGAAGGCGAMALAFHPDRFGFAGSNGFLYPSNTTNG---AIAAGMOQFGVDNMG 198
DB      1 VACPWAGTALNLAHKHPDQFQAGMSGYL---NTTAPGAGTILRAYMLDTGGFNNAM 57
QY      199 WGAPOGGRWKMDHPWVASLLAQNTRVWV-----WSPNPGASDPAAMIGQTAE--AM 250
DB      58 YGSIINPRRPENDPFMMGGILA--NTDVYISAASGLMSPDDGVRVDHRLTGSVLEFVAM 115
QY      251 GNSRMFYNQYRSVSGHGHFDPASGDNGMSNAPOLGAMSG--DIVGA 297
DB      116 TSTRIMEAKAK-LOGLNPTRADYPMYGIHGAQFNSQLERTQGRVLDVMA 164

RESULT 3
US-11-052-554A-157
; Sequence 157, Application US/11052554A
; Publication No. US20050288866A1
; GENERAL INFORMATION:
; APPLICANT: Sachdeva, et al.
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
; TITLE OF INVENTION: PROTEINS OF THERAPEUTIC POTENTIAL
; FILE REFERENCE: 30853/40359A
; CURRENT APPLICATION NUMBER: US/11/052,554A
; PRIOR FILING DATE: 2005-02-07
; PRIOR APPLICATION NUMBER: US 60/589,227
; PRIOR FILING DATE: 2004-07-20
; PRIOR APPLICATION NUMBER: IN 173/DEL/2004
; PRIOR FILING DATE: 2004-02-06
; NUMBER OF SEQ ID NOS: 763
; SOFTWARE: PatentIn version 3.3
```

```

; SEQ ID NO 157
; LENGTH: 783
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis H37Rv
; US-11-052-554A-157

Query Match          7.5%; Score 121.5; DB 7; Length 783;
Best Local Similarity 23.2%; Pred. No. 0.027;
Matches 72; Conservative 32; Mismatches 135; Indels 71; Gaps 13;

QY      4 RSALPALMTIALSPFGLGVAAAEPTAKAPYEINLWVSPSGRDIPIVAF-----LAG 57
DB      68 QAAVAAGVFEVALSAGAGAYH-----SAPAAAGAVLANRAGSVQDILLAAVNAAGYALTG 122
QY      58 GPHAVYILDAFYNAPDV-----SNWVTAGNANNTLAGISVAVAGAGAYSMYTWEDD 111
DB      123 RP---LIGNGANGAPGTGANGAPCGWLLANGGAGSAAAGSLPGAGGAAGLFTGAGAG 179
QY      112 GSKQMDTFLSAEI-----PDWLANRGLAPGHAAVGAAGGYGAMALAAFHDPFGFA 165
DB      180 GAGGSSVVDGDEAGAGAGSGSWLLGTGV--GGVGGIAGAGAGAGGV-----GGA 227
QY      166 GSWSGFLYPSNTTNGAIAAGMOQFGVDNMGWGAPOGRWKMDHPVYASLLAQNTR 225
DB      228 GGLGAGGAGAGAGGLGAVTGV---CG--TGAGG-----LLAGILA----- 264
QY      226 VWVWSPTNPGASDPAAMI--GQTAEMGNSRMFYNGRVSCHGHFDPASGDNGMSW 283
DB      265 ----GPGAGAGTGRGFLNNGVGAGAGNAGLLF---GAGGTGSGSAGAGLGGDGAAGA 316
QY      284 APOLGAMSGD 293
DB      317 GGNTGVLFGN 326

RESULT 4
US-11-188-298-11967
; Sequence 11967, Application US/11188298
; Publication No. US2006007522A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: GENES AND USES FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-21(53452)B
; CURRENT APPLICATION NUMBER: US/11/188,298
; PRIOR FILING DATE: 2005-07-22
; PRIOR APPLICATION NUMBER: 60/592,978
; PRIOR FILING DATE: 2004-07-31
; NUMBER OF SEQ ID NOS: 22569
; SEQ ID NO 11967
; LENGTH: 388
; TYPE: PRT
; ORGANISM: Synecchococcus sp. WH 8103
; US-11-188-298-11967

Query Match          6.7%; Score 108.5; DB 7; Length 388;
Best Local Similarity 22.3%; Pred. No. 0.14;
Matches 79; Conservative 41; Mismatches 113; Indels 121; Gaps 19;

QY      1 MKGRSALLRLAMIALSPFGLGVAAAEPTAKAPYEINLWVSPSGRDIPIVAFLAGSPH 60
DB      31 IRGFSNNRMTWLAACPALMGIGITF-LSAKA---BEL-----PDLANAFLANN-- 76
QY      61 AVAILDA-----FNAGPDVSNVNTAG-----NANNTLAGKISVAPAGAY-----S 103
DB      77 -LMLVATIVITMNG-----FANVEAGMCRQKNAVILN-KNLFVALAVTAVFVGYSP 131
QY      104 MY-----TWEDQSKQMDTFLSAEI-PDWLANRGLAPGHAAVGAAGGYGAMALAAFH 158
DB      132 MYGDAIDGWLFGGFFDPPTVAE-----TISDAGVLPYVDLFQAAFGTAATIVSGIV 187
QY      159 PDR--RG-----FAGSNGFLYPSNTTNGAIAAGMOQFGVDNMGWGAPOGRWKMDHP 212
DB      188 AERIKFGEVVFALITLAFY-----FVAGSWENNG 219
```



```

Db      231 -----NANLGHV-----NFGSGNFGNNGISASIGKNNIG-----FQNLGSNNV 269
QY      170 --GFLPSTTTTNGAIAAGQOFGVDTN-----GMMGAPOIGRWKMDPWHASLLAQ 221
Db      270 GUGNLNLNLT--GFRNTGNGFNGTGNNTNIGIGITGNQIG-----310
QY      222 NNTRVWWSPTNPASDPAAMIGOTAEAMGNSRMFYNQYRSVG--GHNGHFDPPASGDV 278
Db      311 -----IGLNSGTGNFGLFNSSGNGVGFNNGNFGIGNSGNF 349
QY      279 GMGSW 283
Db      350 NTGGM 354

RESULT 8
US-11-188-298-11858
; Sequence 11858, Application US/11188298
; Publication No. US20060075522A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: GENES AND USES FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-21(53452)B
; CURRENT APPLICATION NUMBER: US/11/188,298
; CURRENT FILING DATE: 2005-07-22
; PRIOR APPLICATION NUMBER: 60/592,978
; PRIOR FILING DATE: 2004-07-31
; NUMBER OF SEQ ID NOS: 22569
; SEQ ID NO 11858
; LENGTH: 492
; TYPE: PRT
; ORGANISM: Prochlorococcus marinus str. MIT 9313
US-11-188-298-11858

Query Match      6.6%; Score 106.5; DB 7; Length 492;
Best Local Similarity 21.2%; Pred. No. 0.28;
Matches 77; Conservative 45; Mismatches 114; Indels 127; Gaps 20;

QY      1 MKRSLRLRL-----WTAALSFGLGVAAVEPTAKAPYENLWPPSMGRDI 50
Db      22 LRGPMLLRLRSIRGSSNRSLTWLACVPLALFGLGF-----NLSAHAAEM--PEL 69
QY      51 PVAFLAGPRAVYLLDA-----FNAGPDVSNWVTAG-----NAMNTIAGKGISVVAAPAG 99
Db      70 NAFLFLANN--LWLLVATLIVIFNNAG--FAWVEAEMCGKQKAVNLLA-KNLFVFPALAV 122
QY      100 GAY-----SMY-----TNWBQDSKOMDTFLSAELPDWTLAANKGLAPGCHAAVGAAGG 148
Db      123 SSYWPFGYSIMYGDPPVSAGWLYFNGLFFDPDAVTPB-----LISEAGLVPSVDPLFQAFAAG 178
QY      149 YGAMALAAHPDR--FG-----PAGSMGSPLYPSVTTTNGAIAAGMQOFGVDTNMGWAP 202
Db      179 TAATTVSGLAVERKFEFVPSLVLTGFTYF-----IA-----212
QY      203 QLGRWKMDPWHASL-----LAQNTT--RYWVWSPTNPASDPAAMIGOTAEAMGN 252
Db      213 --GSWENNGMLNTAPEGVEFIDPFAGSIVHSVGA-----AGLVAMTLGPRI 260
QY      253 SRMFYQYRSVGGHN-----GHFDPASDNGKSWAPOLG-----AMGDI 294
Db      261 GKRVGKAQAI PGHNMGIATIGALITWIGVGFNPSQLAMDQVFPYAVATTTTAAAGGA 320
QY      295 VGA 297
Db      321 IGA 323

RESULT 9
US-11-052-554A-172
; Sequence 172, Application US/11052554A
; Publication No. US2005028866A1
; GENERAL INFORMATION:

```

```

; APPLICANT: Sachdeva, et al.
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
; FILE REFERENCE: 30853/40359A
; CURRENT APPLICATION NUMBER: US/11/052,554A
; PRIOR APPLICATION NUMBER: US 60/589,227
; PRIOR FILING DATE: 2004-07-20
; PRIOR APPLICATION NUMBER: IN 173/DEL/2004
; PRIOR FILING DATE: 2004-02-06
; NUMBER OF SEQ ID NOS: 763
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 172
; LENGTH: 1643
; TYPE: PRT
; ORGANISM: Rickettsia prowazekii
US-11-052-554A-172

Query Match      6.6%; Score 106.5; DB 7; Length 1643;
Best Local Similarity 24.7%; Pred. No. 1.1;
Matches 71; Conservative 36; Mismatches 134; Indels 47; Gaps 14;

QY      19 GIGGVAA-----EPTAKAPY---ENLWVPSPMGR-DIPVAFIA-----CG 58
Db      170 GLGSITPAAAPSVLENNLNPITQAPLTIGANSKIVNGNGTNTTNGFIQVSDNTFPAG 229
QY      59 PHAVYLLDA-----FNAGPDVSNWV---TAGNAN--NTLAGKGISVVAAPAGAYSMTYWE 109
Db      230 IKTNIDDOGLMFNSTPDAANTLANLQVGNTINFNGIDGTGLVLSKNGA--ATEFN 286
QY      110 QDSKOMDTFLSAELPDWTLAANKGLAPGGA-AVGAAGGYGAMALAAHPDRPFRAGSN 168
Db      287 VTGTLCNKGKILTELTAAVAGGLISGGAAANAVIGTDGAGAPAFIVSDN-GNAAT 345
QY      169 SGFLPSTTTTNGAIAAGQOFGVDTNMGWAPOLGRWKMDPWHASLLAQNTRVWV 228
Db      346 SGGVYAKNMYVIGSANGGVTFEHIYDVIGLGT--NRTADSKY---IITNSN-----395
QY      229 WSPNPGASDPAAMIGOTAEAMGNSRMFYNQYRSVGHNGHFDPPASG 276
Db      396 FGSTNGNLDTOIVVPDTIKGN---FIGDVKNNGNTAGVITFNANG 440

RESULT 10
US-11-052-554A-162
; Sequence 162, Application US/11052554A
; Publication No. US2005028866A1
; GENERAL INFORMATION:
; APPLICANT: Sachdeva, et al.
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
; FILE REFERENCE: 30853/40359A
; CURRENT APPLICATION NUMBER: US/11/052,554A
; PRIOR APPLICATION NUMBER: US 60/589,227
; PRIOR FILING DATE: 2004-07-20
; PRIOR APPLICATION NUMBER: IN 173/DEL/2004
; PRIOR FILING DATE: 2004-02-06
; NUMBER OF SEQ ID NOS: 763
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 162
; LENGTH: 543
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis H37Rv
US-11-052-554A-162

Query Match      6.5%; Score 106; DB 7; Length 543;
Best Local Similarity 24.1%; Pred. No. 0.34;
Matches 72; Conservative 24; Mismatches 121; Indels 82; Gaps 12;

QY      16 LSFGLGVAAVEPTAKAPYENLWPPSMGRDIVAFIAGGPHAVYLLDAFNAGPDVS 75
Db      280 LFLGLGGDGGAGG-----TSNNNGDGGAGGTTAG-----RLPFLGGDGG 319

```

OY 76 NWVTAGNANTLACKGISVVAAPAGAYSMYTNMBODSKOMDTLSAELPDMILANRGLA 135
| : : : : :
DB 320 NG-GAGTAIGSNAGDG-----GAGDSSALIGYAQSSGGLGSP-----GSTGGDGLG 368
| : : : : :
OY 136 PGHAIVGAAGGAGMALAAPHDRFGFAGSMGFLYPSNTTNGAIAAGMOQFGVDVT 195
| : : : : :
DB 369 GAGAVLIGTVGVGGFGLGGS---NGTGAAGAGG---TGATLIGLAGGGGGIGGPAV 421
| : : : : :
OY 196 N-----GMMGAPDLGRKWKHDPMVHASLLAQNNTVRVWMSPTNPGASDPAAMITGOTABA 249
| : : : : :
DB 422 NVGNVGGLGGGGGG-----NALIGL-----GAGAGAGAGATVVG 458
| : : : : :
OY 250 MANS-----RMFTYNQYRSVGGHNGHFPDPASGDNGM-----GSWAPOLGAMSGD 293
| : : : : :
DB 459 LGGNGGDKGDDGGGLFSTGVGGDGNAGAMPANAGNAGNAGNAPSPSPVFGPGN 517
| : : : : :
RESULT 11
US-11-172-740-542
; Sequence 542, Application US/11172740
; Publication No. US20060057724A1
; GENERAL INFORMATION:
; APPLICANT: MASCIA, Peter
; APPLICANT: ALEXANDROV, Nikolai
; APPLICANT: BROVER, Vyacheslav
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES AND POLYPEPTIDES ENCODED THEREBY USEFUL FOR
; TITLE OF INVENTION: PLANT CHARACTERISTICS AND PHENOTYPES
; FILE REFERENCE: 2750-1602PUS2
; CURRENT APPLICATION NUMBER: US/11/172,740
; PRIOR FILING DATE: 2005-06-30
; PRIOR APPLICATION NUMBER: 60/583,621
; PRIOR FILING DATE: 2004-06-30
; PRIOR APPLICATION NUMBER: 60/584,829
; PRIOR FILING DATE: 2004-06-30
; PRIOR APPLICATION NUMBER: 60/584,800
; PRIOR FILING DATE: 2004-06-30
; NUMBER OF SEQ ID NOS: 2523
; SEQ ID NO 542
; LENGTH: 292
; TYPE: PRT
; ORGANISM: Trifolium aestivum
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)-(292)
; OTHER INFORMATION: Public GI no. 57900400
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION:
; OTHER INFORMATION: Utility: Useful for delaying flowering time
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION:
; OTHER INFORMATION: Utility: Useful for increasing chlorophyll and photosynthetic cap
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION:
; OTHER INFORMATION: Utility: Useful for making ornamental plants with modified leaves
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION:
; OTHER INFORMATION: Utility: Useful for making plants with altered leaf shape eg curl
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION:
; OTHER INFORMATION: Utility: Useful for making plants with increased biomass and foli
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION:

; OTHER INFORMATION: Utility: Useful for making shorter plants and plants with modifi
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION:
; OTHER INFORMATION: Utility: Useful for making smaller plants
US-11-172-740-542
Query Match 6.5%; Score 105.5; DB 7; Length 292;
Best Local Similarity 21.9%; Pred. No. 0.18;
Matches 43; Conservative 22; Mismatches 72; Indels 59; Gaps 7;
OY 44 PSMGRDIPVAFPLAGGPHAVTLDAFNAGPVSMTNNTAGNANTLACKGISVVAAP----- 97
| : : : : :
DB 49 PSPASDLPV-----LYTSLGLTCNDE--NFVTGAGQRAAAAGGIALVAADTSPRG 97
| : : : : :
OY 98 -----AGAYSMYTNMBODSKOMDTLSAELPDMILANRGLAAGGHAAY-G 143
| : : : : :
DB 98 LNTBGAHSWDFGVGAGFYLNATNKKKRNMTDYVKELPKVLSDFBOLNTSRASIFG 157
| : : : : :
OY 144 AAGGYGAMALAAFPDRFGFAGSMGFLYPSNTTNGAIAAGMOQFGVDYNGMG--- 200
| : : : : :
DB 158 HSMGGRGALITIKNTDKYSVASPVPVNPICP-----WGQKA 197
| : : : : :
OY 201 -----APOLGRWKMD 211
| : : : : :
DB 198 FSNYLGPASKDMKEYD 213
| : : : : :
RESULT 12
US-11-188-298-17488
; Sequence 17488, Application US/11188298
; Publication No. US20060075522A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: GENES AND USES FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-21(53452)B
; CURRENT APPLICATION NUMBER: US/11/188,298
; PRIOR FILING DATE: 2005-07-22
; PRIOR APPLICATION NUMBER: 60/592,978
; PRIOR FILING DATE: 2004-07-31
; NUMBER OF SEQ ID NOS: 22569
; SEQ ID NO 17488
; LENGTH: 486
; TYPE: PRT
; ORGANISM: Prochlorococcus marinus subsp. pastoris str. CCMPI378
US-11-188-298-17488
Query Match 6.5%; Score 105.5; DB 7; Length 486;
Best Local Similarity 25.1%; Pred. No. 0.33;
Matches 72; Conservative 39; Mismatches 99; Indels 77; Gaps 19;
OY 1 MKGRSALRLMIALAISFGIGVAVAAEPFAKAPYENLMVPPSPMGRDIP---VAFLAG 57
| : : : : :
DB 32 IRGFSNSRSLMTATVPLALFGLGI-FNLAAHA-----DLBELNAFLAN 76
| : : : : :
OY 58 GPHAVYLLDA-----FNAGPDVSNWYTAG-----NAMNTIAGKISVVAAPAGAY----- 102
| : : : : :
DB 77 N---LWLLATITIVTMNAG---FAMVBAGMCKSKKAVNILA-KULFVPLAATVSYFIG 129
| : : : : :
OY 103 -----SMYTNMBODSKOMDTLSAELPDMILANRGLAAGGHAAGGAGMALA 155
| : : : : :
DB 130 YSLMWGSSVADGWLFGGLFDFPTVAD---WVTADGLVPTVDLFGSAPAGTAATVVS 185
| : : : : :
OY 156 AHPDR-FG---FAGSMGFLYPS-NTTNGA--IAAGMOQFGS---VDNNGMG-- 200
| : : : : :
DB 186 GLVAERKGEFVVFVILVLAFTYPIAGSKWKMGWMLDSIGFVDFAGSSIVHSGAVAGL 245
| : : : : :
OY 201 -----APOLGRWKMDPMV---HASLLAQNNTVRV-WSPNPG 236
| : : : : :
DB 246 VGMMLGPRIGKSDGKPMQPMGHMAIALTGLIIMIGYGRNPS 292
| : : : : :
RESULT 13

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 14, 2006, 17:03:40 ; Search time 88.341 Seconds
(without alignments)
1133.996 Million cell updates/sec

Title: US-10-620-246-66

Perfect score: 1192

Sequence: 1 MALEYSLVLRFTDSDGNFGN.....PSGWRVAGRVSDGVAQLD 228

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_21:.*
1: geneseqp1908:.*
2: geneseqp1908:.*
3: geneseqp2008:.*
4: geneseqp2008:.*
5: geneseqp20029:.*
6: geneseqp20038:.*
7: geneseqp20038:.*
8: geneseqp20048:.*
9: geneseqp20058:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1192	100.0	228	2	AAW72909
2	1192	100.0	228	2	AAW72909
3	159.5	13.4	280	4	AAW72909
4	153.5	12.9	292	6	ADN33271
5	151.5	12.7	278	8	ADN33271
6	145.5	12.2	278	2	AAW72909
7	145.5	12.2	278	2	AAW72909
8	145.5	12.2	278	2	AAW72909
9	145.5	12.2	278	2	AAW72909
10	143.5	12.0	211	4	AAW72909
11	143.5	12.0	302	3	AAW72909
12	129	10.8	331	7	ABO73240
13	127.5	10.7	273	6	ABM67317
14	125	10.5	270	7	ABO72188
15	118	9.9	335	7	ABO78267
16	115	9.6	295	7	ADP06706
17	112.5	9.4	274	6	ADP06706
18	110.5	9.3	106	5	ABP04085
19	110	9.2	288	4	AAW78698
20	110	9.2	375	8	ADG30600
21	110	9.2	375	4	ABG32460
22	100.5	8.4	255	9	AAW04518
23	100.5	8.4	313	7	ABO71801
24	100.5	8.4	485	4	AAU54108

25	100.5	8.4	485	6	ABM50627
26	99	8.3	217	4	ABG22462
27	99	8.3	284	8	ADG44848
28	97.5	8.2	282	5	ABR48896
29	96.5	8.1	513	6	ABU41093
30	96.5	8.1	518	7	ADP06351
31	95	8.0	1107	7	ABU31207
32	94.5	7.9	271	6	ADP04061
33	94	7.9	279	8	ADG22855
34	93	7.8	336	7	ABO73278
35	92.5	7.8	3808	8	ADK16029
36	92.5	7.8	5674	8	ADU47676
37	92.5	7.8	6043	8	ADU47673
38	92	7.7	281	8	ADG44848
39	92	7.7	315	8	ADK78415
40	92	7.7	315	8	ADK78415
41	92	7.7	337	8	ADG33140
42	92	7.7	1482	3	AAW84727
43	92	7.7	1488	3	AAW84725
44	92	7.7	1509	3	AAW84728
45	92	7.7	1517	3	AAW84726

ALIGNMENTS

RESULT 1	AAW72909	standard; protein; 228 AA.
ID	AAW72909	
AC	AAW72909;	
DT	21-JAN-1999	(first entry)
DE	Mycobacterium tuberculosis antigen CFP25A.	
KW	Mycobacterium tuberculosis; antigen; vaccine; immunological; immunogen; infection.	
OS	Mycobacterium tuberculosis.	
PN	W09844119-A1.	
PD	08-OCT-1998.	
PF	01-APR-1998;	98MO-DK000132.
PR	02-APR-1997;	97DK-00000376.
PR	18-APR-1997;	97US-0044624P.
PR	10-NOV-1997;	97DK-00001277.
PR	05-JAN-1998;	98US-0070488P.
XX	(STAT-) STAINS SERUM INST.	
PA	Andersen P, Nielsen R, Rosenkrands I, Wellingh K, Rasmussen PB, Oettinger T, Florio W;	
XX	WPI; 1998-542705/46.	
DR	N-PSDB; AAW63933.	
PT	New isolated mycobacteria polypeptides and nucleic acids - used for developing products for the diagnosis of or vaccination against mycobacterial infections, particularly tuberculosis.	
PS	Claim 1, Page 177-178, 163pp; English.	
XX	The present sequence represents a Mycobacterium tuberculosis protein.	
CC	Products from the present invention, which describes protein fragments and nucleic acid fragments derived from M. tuberculosis, can be used in the detection of and prevention of mycobacterial infections. In particular, the proteins and nucleic acids can be used for the diagnosis of or vaccination against tuberculosis caused by M. tuberculosis, M. africanum or M. bovis	


```

XX      Sequence 228 AA;
SQ
Query Match      100.0%; Score 1192; DB 2; Length 228;
Best Local Similarity 100.0%; Pred. No. 8, 8e-114;
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MAIEVSVLAVFTDSDGNFNGPLGVINASKVHRDRQOLAAQSGYSEITFVDLPSPGSTTA 60
DB      1 MAIEVSVLAVFTDSDGNFNGPLGVINASKVHRDRQOLAAQSGYSEITFVDLPSPGSTTA 60
QY      61 HATHTPRTEIPPAHGPTVGASWMLERGTPIINTLOVPAGIVQVSYHGDLTAISASEWA 120
DB      61 HATHTPRTEIPPAHGPTVGASWMLERGTPIINTLOVPAGIVQVSYHGDLTAISASEWA 120
QY      121 PEFAIHDLSDLDLAAADPADPPDDIAHYLMTWTDTSAGSLRARMFANLGVTEDEATGA 180
DB      121 PEFAIHDLSDLDLAAADPADPPDDIAHYLMTWTDTSAGSLRARMFANLGVTEDEATGA 180
QY      181 AAIRITDYLSDRLTTIQGKSLIHTTWSPEGWVRVAGRVVSDGVAQLD 228
DB      181 AAIRITDYLSDRLTTIQGKSLIHTTWSPEGWVRVAGRVVSDGVAQLD 228

RESULT 2
AA21926 standard; protein; 228 AA.
XX      AA21926;
XX      06-SEP-1999 (first entry)
XX
DE      Amino acid sequence of antigen CFP25A.
XX
KW      Immunogenic; Mycobacterium tuberculosis; immune response; infection;
KW      tuberculosis; fusion polypeptide; T-cell epitope; BSAT-6; MPT59; TB;
KW      pharmaceutical; vaccination; M. africanum; M. bovis; CFP7A; CFP30A;
KW      CFP7B; CFP19; CFP27; CFP30A; RDI-ORF; CFP10A; CFP16; CFP19; CFP23;
KW      CFP25A; CFP30B; CFP7B.
XX
OS      Mycobacterium tuberculosis.
XX
PN      WO924577-A1.
XX
PD      20-MAY-1999.
XX
PF      08-OCT-1998; 98WO-DK000438.
XX
PR      10-NOV-1997; 97DK-00001277.
PR      05-JAN-1998; 98US-0070488P.
PR      01-APR-1998; 98WO-DK000132.
XX
PA      (STAT-) STATENS SERUM INST.
XX
PI      Andersen P, Skjot R;
XX
DR      WPI; 1999-347282/29.
XX
PT      N-PSDB; AA81040.
XX
PT      New immunogenic fragment of Mycobacterium tuberculosis.
XX
XX      Example 3; Page 196-197; 265pp; English.
XX
CC      The invention describes a substantially pure immunogenic polypeptide
CC      fragment (I) from Mycobacterium tuberculosis that is able to evoke a
CC      protective immune response against infections by mycobacteria belonging
CC      to the tuberculosis complex. The invention provides a (1) fusion
CC      polypeptide comprising at least one polypeptide fragment (I) and at least
CC      one fusion partner; (2) a fusion polypeptide fragment comprising a T-cell
CC      epitope from M. tuberculosis protein BSAT-6, or MPT59 and a second
CC      different amino acid sequence from M. tuberculosis, and/or including a
CC      sequence which protects the first amino acid sequence from in vivo
CC      degradation or post-translational processing; (3) a nucleic acid fragment

```

```

CC      that encodes the above polypeptides. The polypeptides and nucleic acid
CC      are useful as pharmaceuticals, for diagnosis of and as antigens for
CC      vaccination against TB caused by Mycobacterium tuberculosis, africanum or
CC      bovis. The polypeptides are also useful for diagnosing ongoing or
CC      previous sensitization in an animal with bacteria belonging to the
CC      tuberculosis complex. The invention also describes the use of CFP7A or
CC      CFP30A or a T-cell epitope of for the induction of a strong immune
CC      response in a mammal; use of CFP7B, CFP19 or MPT59-BSAT6 or a T-cell
CC      epitope of for diagnosis of TB in a mammal by performing a DTH type skin
CC      test; use of CFP27, CFP30A, RDI-ORF2, RDI-ORF3, RDI-ORF5, MPT59-BSAT6,
CC      BSAT6-MPT59, CFP10A, CFP16, CFP19, CFP23, CFP25A, CFP30B, CFP7B or a T-
CC      cell epitope of for the preparation of an immunological composition; and
CC      for the preparation of a subunit vaccine
XX
SQ      Sequence 228 AA;
Query Match      100.0%; Score 1192; DB 2; Length 228;
Best Local Similarity 100.0%; Pred. No. 8, 8e-114;
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MAIEVSVLAVFTDSDGNFNGPLGVINASKVHRDRQOLAAQSGYSEITFVDLPSPGSTTA 60
DB      1 MAIEVSVLAVFTDSDGNFNGPLGVINASKVHRDRQOLAAQSGYSEITFVDLPSPGSTTA 60
QY      61 HATHTPRTEIPPAHGPTVGASWMLERGTPIINTLOVPAGIVQVSYHGDLTAISASEWA 120
DB      61 HATHTPRTEIPPAHGPTVGASWMLERGTPIINTLOVPAGIVQVSYHGDLTAISASEWA 120
QY      121 PEFAIHDLSDLDLAAADPADPPDDIAHYLMTWTDTSAGSLRARMFANLGVTEDEATGA 180
DB      121 PEFAIHDLSDLDLAAADPADPPDDIAHYLMTWTDTSAGSLRARMFANLGVTEDEATGA 180
QY      181 AAIRITDYLSDRLTTIQGKSLIHTTWSPEGWVRVAGRVVSDGVAQLD 228
DB      181 AAIRITDYLSDRLTTIQGKSLIHTTWSPEGWVRVAGRVVSDGVAQLD 228

RESULT 3
AA292764 standard; protein; 280 AA.
XX      AA292764;
XX      26-SEP-2001 (first entry)
XX
DE      C glutamicum protein fragment SEQ ID NO: 6518.
XX
KW      Corynebacterium; amino acid synthesis; vitamin; saccharide;
KW      organic acid synthesis.
XX
OS      Corynebacterium glutamicum.
XX
PN      EP1108790-A2.
XX
PD      20-JUN-2001.
XX
PF      18-DEC-2000; 2000EP-00127688.
XX
PR      16-DEC-1999; 99JP-00377484.
PR      07-APR-2000; 2000JP-00159162.
PR      03-AUG-2000; 2000JP-00280988.
XX
PA      (KYOW ) KYOWA HAKKO KOGYO KK.
XX
PI      Nakagawa S, Mizoguchi H, Ando S, Hayaishi M, Ochial K, Yokoi H;
PI      Tateishi N, Senoh A, Iweda M, Ozaki A;
XX
DR      WPI; 2001-376931/40.
XX
PT      N-PSDB; AA67983.
XX
PT      Novel polynucleotides derived from Corynebacterium bacteria, for identifying
PT      mutation point of a gene, measuring expression of a gene, analyzing
PT      expression profile or pattern of a gene and identifying homologous gene.

```


XX claim 17, SEQ ID NO 6518; 246bp + Sequence Listing; English.
 PS The present invention provides a number of nucleotide and protein
 CC sequences from the Corynebacterium bacterium Corynebacterium glutamicum. These
 CC are useful for identifying the mutation point of a gene derived from a
 CC mutant of corynebacterium bacterium, measuring expression amount and analyzing
 CC the expression profile or expression pattern of a gene derived from
 CC Corynebacterium bacterium, and identifying a homologue of a gene derived from
 CC corynebacterium bacterium. Corynebacterium bacteria are useful for producing amino
 CC acids, nucleic acids, vitamins, saccharides and organic acids.
 CC particularly L-lysine. The present sequence is a protein described in the
 CC exemplification of the invention. Note: The sequence data for this patent
 CC did not form part of the printed specification, but was obtained in
 CC electronic format directly from the European Patent Office

XX Sequence 280 AA;

Query Match 13.4%; Score 159.5; DB 4; Length 280;
 Best Local Similarity 25.9%; Pred. No. 1.4e-07;
 Matches 73; Conservative 33; Mismatches 95; Indels 81; Gaps 15;

QY 10 VFTDSQNGNPNPGVY-NMSKVHHRDQQLAAGSGVSETFVDLPSPGTTAATITPR 68
 DB 14 VFS-SSEPMGNPVAIVADADLSAQWARIARWTNLSSETFLKPTQEGADYVRITPT 72
 QY 69 TEIPFAGHPVTVGASWMLR-----ERGTPINTLOVPAGIVOV-SYHGDLTAISSEMAPR 123
 DB 73 GELPFGHPFTLGTAVHAFRELBHGEQGTOL-VQECVAGLVAVRAIDPAGSGLAFQA--PPTL 129
 QY 124 AIHDLDSLALAAADPADP-PDDI-AHYLMTWTRDSNG----- 159
 DB 130 KQGPDLASDLDAACBALGISPDFIRAH---QWVNDNGMAVVELPSAQHYLDLEPDSAH 186
 QY 160 -----SLRPMFPAANTGVTEDEATGAALIRITDYLSRDLTTQG---- 198
 DB 187 PTLKLGVTGAYPEGAPPAFVRAFAQIG--EDPVVSGSLNASIAQWLRHDSRAGEGYLAS 244
 QY 199 KGSLLI-----HTTWSPEGWVRVAGRV--VSDGVAQ 226
 DB 245 QGTAIGRAGRHHISHSIAW-----VGGSVTTTIFGTAH 279

RESULT 4
 ADA33271

ID ADA33271 standard; protein; 292 AA.

AC ADA33271;

DT 20-NOV-2003 (first entry)

DE Acinetobacter baumannii protein #432.

KW Acinetobacter baumannii; bacterial disease; antibacterial; vaccine;

KM plant biocontrol agent.

OS Acinetobacter baumannii.

PN US6562958-B1.

PD 13-MAY-2003.

PF 04-JUN-1999; 99US-00328352.

PR 09-JUN-1998; 98US-0088701P.

PA (GENO-) GENOME THERAPEUTICS CORP.

PI Breton G, Bush D;

DR WPI, 2003-576092/54.

XX N-PSDB; ADA29145.

PT New Acinetobacter baumannii proteins and nucleic acids, useful as reagents
 PT for diagnosing a bacterial disease, as components of antibacterial
 PT vaccines, as targets for antibacterial drugs, or as biocontrol agents for
 PT plants.

XX Example; SEQ ID NO 4558; 328bp; English.

XX The invention relates to isolated Acinetobacter baumannii nucleic acids.
 CC The A. baumannii nucleic acids and polypeptides are useful as reagents
 CC for diagnosing a bacterial disease, as components of antibacterial
 CC vaccines, as targets for antibacterial drugs, to detect the presence of
 CC A. baumannii and other Acinetobacter species in a sample, in screening
 CC compounds for the ability to interfere with the A. baumannii life cycle
 CC or to inhibit A. baumannii infection, and as biocontrol agents for
 CC plants. The present sequence represents the amino acid sequence of an A.
 CC baumannii protein.

XX Sequence 292 AA;

Query Match 12.9%; Score 153.5; DB 6; Length 292;
 Best Local Similarity 25.6%; Pred. No. 6.2e-07;
 Matches 61; Conservative 28; Mismatches 80; Indels 69; Gaps 8;

QY 10 VFTDSQNGNPNPGVY-NMSKVHHRDQQLAAGSGVSETFVDLPSPGTTAATITPR 68
 DB 14 VFT-SQAFKGNPVAIVMDASTLTSEQWQALANTNLSSETFLVLPPTDSQADYQVRITPTQ 72
 QY 69 TEIPFAGHPVTVGASWMLRERG-----TP 91
 DB 73 SELPFGHPFTLGTAVHAFRELBHGEQGTOL-VQECVAGLVAVRAIDPAGSGLAFQA--PPTL 132
 QY 92 INTLOVP--AGI-----VOVSYHGDLTAISARSEMAEFALHDLDSLALAAADPADPDD 145
 DB 133 LDALQTEKALIELKCKINQWNAALVDGAR--WVVLQAVN-----AKAVLASQPD 182
 QY 146 IAHYLMTWTRDSAGS-----LRPMFPAANTGVTEDEATGAALIRITDYLSRDLTTQG---- 189
 DB 183 NALKQSLDMKVTGTYTGYFERSNEQKRIEVRSPAYGVNEDPVGSGNGSVASPI 240

RESULT 5
 ADS15066

ID ADS15066 standard; protein; 278 AA.

AC ADS15066;

DT 16-DEC-2004 (first entry)

DE Pseudomonas aeruginosa quorum sensing controlled protein, SEQ ID 621.

KW Cytostatic; Anti-HIV; Antimicrobial; Immunosuppressive;

KM quorum sensing signaling; bacterium; quorum sensing controlled gene;

KW biofilm-associated disease; cystic fibrosis; AIDS; middle ear infection;

XX acne; periodontal disease.

OS Pseudomonas aeruginosa.

PN W02004083385-A2.

PD 30-SEP-2004.

PF 11-MAR-2004; 2004WO-US007467.

PR 14-MAR-2003; 2003US-00389647.

PA (IOWA) UNIV IOWA RES FOUND.

PI Greenberg EP, Schuster M, Lostrich C;

DR WPI, 2004-709932/69.

PT Identifying a modulator of quorum sensing signaling in bacteria, useful
 PT for treating a biofilm-associated disorder, comprises contacting the cell

PT with a quorum sensing signal molecule in the presence and absence of a
PT test compound.
XX
XX
XX
XX
XX
PS Disclosure; SEQ ID NO 621; 233pp; English.
CC The present invention relates to a method for identifying a modulator of
CC quorum sensing signalling in bacteria. The method comprises: providing a
CC cell that comprises a quorum sensing controlled gene (ADS1446-ADS14798),
CC where the cell is responsive to a quorum sensing signal molecule such
CC that a detectable signal is generated; contacting the cell with a quorum
CC sensing signal molecule in the presence and absence of a test compound;
CC and detecting a change in the detectable signal. The method and modulator
CC identified by the method are useful for treating a biotilm-associated
CC disease or disorder, e.g. cystic fibrosis, AIDS, middle ear infections,
CC acne, periodontal disease, catheter-associated infections, and medical
CC device-associated infections. The present sequence is a quorum sensing
CC controlled protein, used to illustrate the invention. Note: The sequence
CC data for this patent was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences
XX
XX
SQ Sequence 278 AA;
Query Match 12.7%; Score 151.5; DB 8; Length 278;
Best Local Similarity 28.1%; Pred. No. 9.3e-07;
Matches 62; Conservative 26; Mismatches 88; Indels 45; Gaps 9;
QY 19 GNPILGV-INASKVHHRDQOLAAGSGSEIFVLDLPSGSTTAHTIHPRTIIPRAGHP 77
DB 17 GNPVAVFPDODDLSGSRMQRARENLSESTFVLRRPQDG-DARIRIFTVNLLPFGHP 75
QY 78 TVGASWML-REKGTPIINTLOVPAGIVQVSYH---GDLTAISARS---EW----- 119
DB 76 LIGTAIALGAEITKDLFLETBMGTVPFALERODGKVAACSMQCP1PTWHEFSRPABLLA 135
QY 120 -----APERAHIDLSDLAADP-----ADPEPDIAHYLMTWDRSA 158
DB 136 ALGLKSTPEIEVRNGPRHVFGLSEVALSLHPDRHALCDP-DLAVNCFAGAGR-- 192
QY 159 GSILRAMPANLGVTEDEATGAARITDYLSRDLTITQK 199
DB 193 -HKRSMPSPAYGVVEDATGSAAGPLAHLAHRRQIPYQ 232
RESULT 6
AAR87535
ID AAR87535 standard; protein; 278 AA.
XX
XX
AC AAR87535;
XX
DT 16-OCT-2003 (revised)
DT 15-OCT-1995 (first entry)
XX
XX
DE Pseudomonas aureofaciens phenazine gene cluster ORF3 protein phz3.
XX
XX
KW Antipathogenic substance; phenazine; antibiotic; fungicide; pesticide.
XX
XX
OS Pseudomonas chlororaphis.
XX
XX
FH Key Location/Qualifiers
FT Misc-difference 210 /note= "unknown"
FT
XX
XX
PN WO9533818-A2.
XX
XX
PD 14-DEC-1995.
XX
XX
PF 30-MAY-1995; 95WO-IB000414.
XX
XX
PR 08-JUN-1994; 94US-00258261.
XX
XX
PA (CIBA) CIBA GEIGY AG.
XX
XX
PI Ligon J, Schupp T, Beck JU, Hill DS, Ryals JA, Gaffney TD;

PI Lam ST, Hammer PB, Uknes SJ;
XX
XX
XX WPI; 1996-040226/04.
DR N-PSDB; AAT06770.
XX
XX
PT New genes for biosynthesis of anti-pathogenic substances - pref.
PT Pyriroclitrin and soraphen, useful for disease control in plants.
XX
XX
PS Disclosure; Page 163-169; 190pp; English.
XX
XX
CC This is the protein phz3 encoded by the phenazine gene cluster ORF3. The
CC gene cluster may be expressed recombinantly to produce phenazine, or 16-
CC OCT-2003 to standardise OS field)
XX
XX
SQ Sequence 278 AA;
Query Match 12.2%; Score 145.5; DB 2; Length 278;
Best Local Similarity 28.3%; Pred. No. 3.9e-06;
Matches 62; Conservative 33; Mismatches 83; Indels 41; Gaps 9;
QY 19 GNPILGV-INASKVHHRDQOLAAGSGSEIFVLDLPSGSTTAHTIHPRTIIPRAGHP 77
DB 17 GNPVAVFPDODDLSAQMQRIRARENLSESTFV-LKPRNCGDALLIFTVNELPFGHP 75
QY 78 TVGASWMLREKGTPIINTLOVPAGIVQVSYH---HGDLTAISARSEWAEFAL-HDLDSL 131
DB 76 LIGTDIALGAR-TDNRRLFLETQMGTIAPFALERONGSVIAASWDQPIPTWTALGRDAEYL 134
QY 132 DALAADPADFPDDIYH-----YLMWTTDRS---AGS---- 160
DB 135 KALGISD-STFPIEIHNGPRHVFGLPSIALSLHPDRHALYSFHDMAINCFAAGRR 193
QY 161 LPARMPANLGVTEDEATGAARITDYLSRDLTITQK 199
DB 194 WSRMSPSPAYGVVEDATGSAAGPLAHLAHRRQIPYQ 232
RESULT 7
AAW31306
ID AAW31306 standard; protein; 278 AA.
XX
XX
AC AAW31306;
XX
DT 17-OCT-2003 (revised)
DT 12-MAR-1998 (first entry)
XX
XX
DE Pseudomonas aureofaciens phenazine phz3 protein.
XX
XX
KW Phenazine; biosynthesis; antibiotic; antipathogenic; transgenic plant;
KW phytopathogen; resistance.
XX
XX
OS Pseudomonas chlororaphis.
XX
XX
FH Key Location/Qualifiers
FT Protein 1..278
FT /label= ORF3
FT /product= "phz3"
FT Misc-difference 210 /label= unknown
FT
XX
XX
PN US5662898-A.
XX
XX
PD 02-SEP-1997.
XX
XX
PF 01-JUN-1995; 95US-00457342.
XX
XX
XX 20-AUG-1990; 90US-00570184.
XX
XX 02-JUL-1992; 92US-00908284.
XX
XX 31-AUG-1992; 92US-00937648.
XX
XX 01-JUL-1993; 93US-00087636.
XX
XX 08-JUN-1994; 94US-00258261.
XX

PA (CIBA) CIBA GEIGY CORP.
 XX Ligon JM, Hill DS, Lam ST, Hammer PB;
 XX WPI; 1997-447901/41.
 DR N-PSDB; AAT89957.
 XX
 PT Protecting plants against pathogens with genetically transformed
 PT biological control agent - which expresses all polypeptide(s) involved in
 PT pyrrolnitrin biosynthetic pathway.
 XX
 PS Example 18; Col 149-150; 86pp; English.
 XX
 CC Proteins AAW31304-W31307 represent the open reading frames (ORF's) 1-4
 CC encoded by the gene cluster phz1, phz2, phz3 and phz4 involved in
 CC phenazine biosynthesis. Phenazines are nitrogen-containing heterocyclic
 CC compounds with a common planar aromatic tricyclic structure. It has been
 CC proposed that phenazine antibiotic function arises from the formation of
 CC intercalative complexes with DNA interfering with DNA metabolism.
 CC Transgenic plants containing such antipathogenic genes should have
 CC enhanced resistance to attack by phytopathogens. (Updated on 17-OCT-2003
 CC to standardise OS field)
 XX
 SQ Sequence 278 AA;
 XX
 Query Match 12.2%; Score 145.5; DB 2; Length 278;
 Best Local Similarity 28.3%; Pred. No. 3.9e-06;
 Matches 62; Conservative 33; Mismatches 83; Indels 41; Gaps 9;
 QY 19 GNPGLV-INSKVEHNDQQLAAGSYSETIPVDLPSPGTTAHTTPETEPFAGHP 77
 DB 17 GNFVAVFDDADLSABQMGRIAREMNSTETTFV-LKPRNGCDALIRIFTVNEHLPFAGHP 75
 QY 78 TVGASWTLRERGTPINTLOVPAGIVQSY-----HGDLTISARSEAPPEAF-HDDLSL 131
 DB 76 LIGTDIALGAR-TDNHRLPLETQMGTTAFELERONGSVIAASMDQPIPTWTALGRDDEL 134
 QY 132 DALAADPADPDDIAH-----YLTWTDRS---AGS--- 160
 DB 135 KALGISD-STFPEIETHNGRHHVFGLPSTIALSALHPDRHALYSFHDMAINCAGARR 193
 QY 161 LRAARMPANIGVTEDEATGAARITDYLSRDLLTTTQK 199
 DB 194 WRSRMFSPAYGVVEDAXTGSAAGPLAIHLARHQIIEFGQ 232
 RESULT 8
 ID AAW69403 standard; protein; 278 AA.
 AC AAW69403;
 XX
 DT 17-OCT-2003 (revised)
 DT 07-DEC-1998 (first entry)
 DE Phenazine gene cluster protein phz3.
 XX
 KM Pyrrolnitrin; biosynthetic pathway; pathogen protection; phenazine;
 KM plant antipathogenic substance production; anti-fungal antibiotic;
 KM fungal respiratory electron transport inhibitor; lipoprotein damage.
 XX
 OS Pseudomonas chlororaphis.
 XX
 PH Key Location/Qualifiers
 FT Misc-difference 210 /note= "unspecified amino acid"
 XX
 XX US5817502-A.
 XX
 PD 06-OCT-1998.
 XX
 PF 09-OCT-1996; 96US-00729214.
 XX

PR 08-JUN-1994; 94US-00258261.
 PR 30-MAY-1995; 95WO-1B000414.
 XX
 PA (NOVS) NOVARTIS FINANCE CORP.
 XX
 PI Kirner S, Hammer PB, Hill DS, Van Pee K, Ligon JM, Lam ST;
 XX WPI; 1998-556391/47.
 DR N-PSDB; AAV58732.
 XX
 PT Genes encoding enzymes of the biosynthetic pathway of pyrrolnitrin -
 PT useful for producing transgenic plants which can produce pyrrolnitrin as
 PT an anti-pathogenic agent.
 XX
 PS Example 18; Col 161-162; 109pp; English.
 XX
 CC This sequence is encoded by the phenazine gene cluster, isolated from
 CC Pseudomonas aureofaciens. This sequence was used to isolate the
 CC pyrrolnitrin gene region of the invention, that encodes at least one
 CC enzyme required in the biosynthetic pathway of pyrrolnitrin. The DNA and
 CC host cells transformed with it are useful for the production of
 CC transgenic plants with protection against phytopathogens. The enzymes are
 CC part of a biosynthetic pathway producing plant antipathogenic substances
 CC (APS). The compound ultimately produced by the pathway, pyrrolnitrin, is
 CC a broad range, anti-fungal antibiotic. It inhibits fungal respiratory
 CC electron transport and causes general lipoprotein damage. The transformed
 CC cells can additionally be used in compositions to be applied to plants to
 CC provide resistance, as can purified APS produced by them. Transgene
 CC dependent resistance eliminates the need to spray crops with chemical
 CC based pesticides and antibiotics, which is expensive and time consuming,
 CC and in addition, especially in the case of antibiotics, their over use
 CC leads to resistance. In addition, transgenic production of these enzymes
 CC overcomes problems of applying micro-organisms which can be slow growing
 CC and isolated in their growth areas. (Updated on 17-OCT-2003 to
 CC standardise OS field)
 XX
 SQ Sequence 278 AA;
 XX
 Query Match 12.2%; Score 145.5; DB 2; Length 278;
 Best Local Similarity 28.3%; Pred. No. 3.9e-06;
 Matches 62; Conservative 33; Mismatches 83; Indels 41; Gaps 9;
 QY 19 GNPGLV-INSKVEHNDQQLAAGSYSETIPVDLPSPGTTAHTTPETEPFAGHP 77
 DB 17 GNFVAVFDDADLSABQMGRIAREMNSTETTFV-LKPRNGCDALIRIFTVNEHLPFAGHP 75
 QY 78 TVGASWTLRERGTPINTLOVPAGIVQSY-----HGDLTISARSEAPPEAF-HDDLSL 131
 DB 76 LIGTDIALGAR-TDNHRLPLETQMGTTAFELERONGSVIAASMDQPIPTWTALGRDDEL 134
 QY 132 DALAADPADPDDIAH-----YLTWTDRS---AGS--- 160
 DB 135 KALGISD-STFPEIETHNGRHHVFGLPSTIALSALHPDRHALYSFHDMAINCAGARR 193
 QY 161 LRAARMPANIGVTEDEATGAARITDYLSRDLLTTTQK 199
 DB 194 WRSRMFSPAYGVVEDAXTGSAAGPLAIHLARHQIIEFGQ 232
 RESULT 9
 ID AAB18558 standard; protein; 278 AA.
 AC AAB18558;
 XX
 DT 12-SEP-2003 (revised)
 DT 15-JAN-2001 (first entry)
 DE Protein phz3 encoded by a phenazine gene region.
 XX
 KM Pyrrolnitrin; PnA, PnB, PnC, PnD; antibiotic; biocontrol;
 KM fungal respiratory electron transport inhibitor; transgenic plant;
 KM antipathogenic substance; biosynthetic gene; fungal resistance;
 XX

CC proteinogenic or nonproteinogenic amino acid (preferred), purine or
 CC pyridine base, nucleoside, nucleotide, lipid, saturated or unsaturated
 CC fatty acid, diol, carbohydrate, aromatic compound, vitamin, cofactor,
 CC polyelectrolyte or enzyme. The amino acids produced can be lysine, glutamine,
 CC glutamate, alanine, aspartate, glycine, serine, threonine, methionine,
 CC cysteine, valine, leucine, isoleucine, arginine, proline, histidine,
 CC tyrosine, phenylalanine, or tryptophan. The fine chemical production can
 CC be modulated. The presence of (1) or HA proteins encoded by then are used
 CC for diagnosing the presence or activity of *Corynebacterium diphtheriae*.
 CC (1) can be used to map the *C. glutamicum* genome or can be used as markers
 CC for genetically engineered *Corynebacterium* or *Brevibacterium*. The HA
 CC proteins encoded by the (1) are used to maintain homeostasis in *C.*
 CC *glutamicum* or help the microorganism to adapt to different environmental
 CC conditions

XX
 XX
 SQ Sequence 211 AA;

Query Match 12.0%; Score 143.5; DB 4; Length 211;
 Best Local Similarity 26.6%; Pred. No. 4.3e-06;
 Matches 58; Conservative 26; Mismatches 79; Indels 55; Gaps 10;

QY 19 GNPLGYI-NASKVEHRDROQLAAGSVSETIFVDLPSPSTTAATHTRTTIPAGHP 77
 DB 2 GNPLAVIADDDISAEQMARIARKTNLSSTTFLLKPTQEGADYKRIPTFGSLPFAQHP 61

QY 78 TVGASWMLR-----ERGTPINTLQVPAGIVQV-SYHGDLTAISASEWAPFAIHDLDSL 132
 DB 62 TLGTAVHFRRLHGGQFQL-VQEGVAGLVAVRAIDGASGAPQA--PFTLKQGPLDASD 118

QY 133 ALAADPADP-PDDI-AHYLTMTDRSAG----- 159
 DB 119 LDAACETALGISPDPIRAH---QWVDNGPGWAVVELPSAQHYLDLEPPSAHPTLKGVIG 175

QY 160 -----SARMPFANLGVTEDEAGAAAIRITDYL 189
 DB 176 AYPGAPHAFVRAFAQIG--EDPVGSLNAFPAQWL 211

RESULT 11

AA95035
 ID AA95035 standard; protein; 302 AA.

XX
 AC AA95035;
 XX
 DT 23-JUN-2000 (first entry)

XX Candida albicans polypeptide sequence # 3.

XX
 KM Candida albicans infection; growth; survival; medicament; AIDS;
 KM vulvovaginitis; immunocompromised patient; treat.

OS Candida albicans.

XX
 PN BP982401-A2.

PD 01-MAR-2000.

PF 23-DEC-1998; 98BP-00310694.

PR 14-AUG-1998; 98GB-00017796.

XX
 PA (JANC) JANSSEN PHARM NV.

PI Contreras RH, Nelissen B, De Backer MD, Luyten WHML, Viaene JF;
 PI Iogshne WG;

XX
 DR WPI; 2000-258614/23.

XX
 PT Essential polypeptides isolated from *Candida albicans*, useful in the
 PT treatment of diseases caused by *C. albicans*, especially in
 PT immunocompromised subjects, e.g., AIDS patients.

XX
 PS Claim 3; Page 19-20; 133p; English.

XX This sequence represents a polypeptide that is critical for the survival
 CC and growth of *Candida albicans*. The *C. albicans* nucleic acid molecules
 CC encoding the polypeptides of the invention may be used as probes and
 CC primers for detecting homologous nucleic acid molecule sequences. The
 CC polypeptides and nucleic acid molecules and compounds identified as
 CC selectively modulating the expression of the polypeptides, may be used as
 CC medicaments or for the preparation of a medicament to treat *C. albicans*
 CC associated diseases, especially in AIDS patients and to treat
 CC vulvovaginitis in otherwise healthy females. The use of the polypeptides
 CC and polynucleotide sequences to treat *C. albicans* associated diseases has
 CC fewer side effects and less toxicity than previously used methods such as
 CC the use of amphotericin. This method is therefore especially suitable for
 CC immunocompromised patients, such as AIDS patients

XX
 XX
 SQ Sequence 302 AA;

Query Match 12.0%; Score 143; DB 3; Length 302;
 Best Local Similarity 24.2%; Pred. No. 7.8e-06;
 Matches 63; Conservative 36; Mismatches 85; Indels 76; Gaps 12;

QY 10 VFTDSQNFNGNPGVI-NASKVEHRDROQLAAGSVSETIFVDLPSPSTTAATHTTP- 67
 DB 9 VFTNVR-YLGNPVAVIDSDNLTQEMQKIAKMTNLSETTIFLTPKSIAXYSRITSG 67

QY 68 RTEIPFAGHPPTVAGSWMLRERG--TPINTLQV---PAGIVQVS----- 105
 DB 68 GNELPFAQHPHTLTGAFALEBDGKIKPNDNQIIOEGCAGLVKISVEKTPNNNSNELPFL 127

QY 106 -----YH--GDLTARISARSEW-----AEPFAIHDLSDIALAADPAD 141
 DB 128 SFELPYEKFHIDKVIIELOHSMGNTNIGKEPVLIDAGKMAVFOIGS-----GKEVLD 182

QY 142 FPDIDIAHYL-----WTWT-----DRSAGSLRMPFANLGVTEDEAGAAAIRITDY 188
 DB 183 LNDYLAQIERLSLENGTTGICVPGKHNEDSVELRIYAVAGVADPACGSGGALGAY 242

QY 189 LSR-----DLTTQGR 199
 DB 243 LANHVFNEKEKFTTIDISQGR 262

RESULT 12

ABO73240
 ID ABO73240 standard; protein; 331 AA.

XX
 AC ABO73240;

XX
 DT 29-JUL-2004 (first entry)

XX Pseudomonas aeruginosa polypeptide #5415.

XX Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.

XX Pseudomonas aeruginosa.

OS Pseudomonas aeruginosa.

XX
 PN US6551795-B1.

PD 22-APR-2003.

PF 18-FEB-1999; 99US-00252991.

PR 18-FEB-1998; 98US-0074788P.

XX
 PA (GENO-) GENOME THERAPEUTICS CORP.

PI Rubenfield MJ, Nolling J, Deloughery C, Bush D;

XX
 DR WPI; 2003-615309/56.

XX
 PT Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,

PT useful as molecular targets for diagnostics, prophylaxis and treatment of
 PT pathological conditions resulting from bacterial infection.

XX Disclosure; SEQ ID NO 21986; 455bp; English.

CC The invention relates to *Pseudomonas aeruginosa* polypeptides and the
 CC polynucleotides encoding them. The sequences are useful in diagnosis and
 CC therapy of pathological conditions, as molecular targets for diagnostics,
 CC prophylaxis and treatment of pathological conditions resulting from a
 CC bacterial infection, for evaluating a compound, such as a polypeptide,
 CC for the ability to bind a *P. aeruginosa* nucleic acid, as components of
 CC effective antibacterial targets, as targets for antibacterial drugs,
 CC including anti-*P. aeruginosa* drugs, as templates for recombinant
 CC production of *P. aeruginosa*-derived peptides or polypeptides, as target
 CC components for diagnosis and/or treatment of *P. aeruginosa*-caused
 CC infection, and in detection of *P. aeruginosa* sequences or other sequences
 CC of *Pseudomonas* species using biotech technology. Sequences AB067826-
 CC AB064396 represent *P. aeruginosa* polypeptides of the invention. Note: The
 CC sequence data for this patent did not form part of the printed
 CC specification but was obtained in electronic format from USPTO at
 CC seqdata.uspto.gov/sequence.html

XX Sequence 331 AA;

Query Match 10.8%; Score 129; DB 7; Length 331;

Best Local Similarity 27.6%; Pred. No. 0.00024;

Matches 51; Conservative 24; Mismatches 78; Indels 32; Gaps 10;

DB 19 GNPGLG-VINASKVEHNDROQLAAGSGYSETFVDPSPSTTNAHTIPRTETPPAGHP 77

DB 85 GNAVAVVIAADRISSKQMLPFAAWTQLSETFPLRTVAADYRVRIFFPLRLPFGHP 144

QY 78 TVGA-SMWLRERG---TPINTLOVPAGIVOVSYHGDULTAISARSEWAPF---AIHD-- 127

DB 145 TIGSCQVWLANOGGNADEIVQECLAGLRIRKGLLSRA-----APRLRGAGVDEY 199

QY 128 LDSLDAALADPADFPDDIAHYLTWTDRSAGSLRARMFANLGVTEDEATGAARITD 187

DB 200 LRRISGLGLSPQVGRS-----QWVDNGPGWVAVRL-----ATRD---VLAIR-PD 243

QY 188 YLSRD 192

DB 244 YAKLD 248

RESULT 13

ABM67317

ID ABM67317 standard; protein; 293 AA.

XX ABM67317;

DT 20-NOV-2003 (first entry)

DE Photorhabdus luminescens protein sequence #414.

KM Antibacterial; fungicide; insecticide; polymorphism; genetic analysis;
 KM detection; food; gene expression; plant; animal; microorganism; toxin;
 KM antibiotic; bioplastic; virulence factor; disease model; plague;
 KM whooping cough.

XX Photorhabdus luminescens.

OS Photorhabdus luminescens.

XX WO200294867-A2.

XX 28-NOV-2002.

XX 07-FEB-2002; 2002WO-IB003040.

XX 07-FEB-2001; 2001FR-00001659.

XX (INSP) INST PASTEUR.

XX (CNRS) CNRS CENT NAT RECH SCI.

PI Duchaud E, Tacurit S, Glaeer P, Frangeul L, Kunst F, Danchin A;
 PI Buchrieser C;
 XX WPI; 2003-148459/14.

PT Genomic sequence of *Photorhabdus luminescens* and encoded polypeptides,
 PT useful e.g. as therapeutic antimicrobials and agricultural pesticides.

XX Claim 2; SEQ ID NO 414; 1205bp; French.

CC The invention relates to the isolation of genes and their encoded
 CC proteins from *Photorhabdus luminescens*. The isolated sequences are
 CC sources of probes and primers for detecting the genome of *P. luminescens*
 CC and related species; to study polymorphisms, for gene analysis and for
 CC detection/amplification of the genes. Antibodies (Ab) raised against the
 CC polypeptides encoded by the genes are used for detection/identification
 CC of *P. luminescens*, e.g. in foods. The genes, proteins, Ab and cells that
 CC carry a gene-containing vector are used to select compounds that
 CC modulate, regulate, induce or inhibit expression of the genes in plants,
 CC animals or microorganisms other than *P. luminescens* and are able to alter
 CC response or sensitivity to toxins and antibiotics produced by *P.*
 CC *luminescens*. Cells transformed to express the genes are useful for
 CC recombinant production of the proteins, particularly toxins and
 CC antibacterials useful as insecticides, bactericides and fungicides. The
 CC genes, proteins, vectors containing the genes and Ab are also useful
 CC therapeutically (to treat microbial infection by bacteria or fungi that
 CC are sensitive to *P. luminescens*-encoded toxins or antibiotics) and as
 CC biopesticides. Other uses of the genes and the proteins are as virulence
 CC factors and for identifying targets of human diseases for which *P.*
 CC *luminescens* is a model (particularly plague and whooping cough). This
 CC sequence represents one of the isolated *P. luminescens* proteins

XX Sequence 293 AA;

Query Match 10.7%; Score 127.5; DB 6; Length 293;

Best Local Similarity 22.7%; Pred. No. 0.00029;

Matches 62; Conservative 39; Mismatches 95; Indels 77; Gaps 13;

QY 19 GNPGLV-NASKVEHNDROQLAAGSGYSETFVDPSPSTTNAH-ATHTPRTETPPAGH 76

DB 19 GNPVAVVMDAGLSSIQMGGINMTNLSETTFL-LEAPNPLADYRVRIFFPLRLPFGHP 77

QY 77 PTVGASWMLRERGTPINTLOVPAG-IVOVSYHGDULTA-ISARSEW-----APRLAHD 127

DB 78 PTGTAAHLLEAG---LQAREGRIVQECAGLTLTNTVERDEGKLTFFELPEPTTP 133

QY 128 LDS-----LDALAADPADFPDDIAH----- 148

DB 134 LSSQIDRLRLSLDCELRALTPALIDVGARMIVAHTGAENVLATKP-DYARLLEHDIQ 192

QY 149 -----YLMTWTDRSAGSLRARMFANLGVTEDEATGAARITDYLRSRLDTTGGKS 201

DB 193 NNITGVCLYGAHHEGADIEVRSFAPSGVNEPVCVSGNSVAAM-FRHHKVAAMIDK 251

QY 202 LHTHT-----W--SPGQVRAVGRVVS 221

DB 252 IVHSQSGKKLGRQGSVWLSHSDGKIFVGSAAVT 284

RESULT 14

ABO72188

ID ABO72188 standard; protein; 270 AA.

XX ABO72188;

DT 29-JUL-2004 (first entry)

DE *Pseudomonas aeruginosa* polypeptide #4363.

XX Bacterial infection; *Pseudomonas aeruginosa* infection; antibacterial.

XX *Pseudomonas aeruginosa*.

PN US6551795-B1.
XX
PD 22-APR-2003.
XX
PF 18-FEB-1999; 99US-00252991.
XX
PR 18-FEB-1998; 98US-0074788P.
XX 27-JUL-1998; 98US-0094190P.
XX
PA (GENO-) GENOME THERAPEUTICS CORP.
PI Rubenfield MJ, Nolling J, Deloughery C, Bush D;
XX WPI; 2003-615309/58.
DR N-PSDB; ABD05759.
XX
PT Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
XX useful as molecular targets for diagnostics, prophylaxis and treatment of
XX pathological conditions resulting from bacterial infection.
XX
PS Disclosure; SEQ ID NO 20934; 455bp; English.
XX
CC The invention relates to Pseudomonas aeruginosa polypeptides and the
CC polynucleotides encoding them. The sequences are useful in diagnosis and
CC therapy of pathological conditions, as molecular targets for diagnostics,
CC prophylaxis and treatment of pathological conditions resulting from a
CC bacterial infection, for evaluating a compound, such as a polypeptide,
CC for the ability to bind a P. aeruginosa nucleic acid, as components of
CC effective antibacterial targets, as targets for antibacterial drugs,
CC including anti-P. aeruginosa drugs, as templates for recombinant
CC production of P. aeruginosa-derived peptides or polypeptides, as target
CC components for diagnosis and/or treatment of P. aeruginosa-caused
CC infection, and in detection of P. aeruginosa sequences or other sequences
CC of Pseudomonas species using biochip technology. Sequences AB067826-
CC AB084396 represent P. aeruginosa polypeptides of the invention. Note: The
CC sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format from USPTO at
CC seqdata.uspto.gov/sequence.html
XX
XX
SQ Sequence 270 AA;
Query Match 10.5%; Score 125; DB 7; Length 270;
Best Local Similarity 26.4%; Pred. No. 0.0047;
Matches 70; Conservative 33; Mismatches 108; Indels 54; Gaps 13;
QY 3 IEVSVARVPTDSQNF-GNPLGVYNASK-VEHNRQQLAASGSEITFPVLDSPGSGTGA 60
DB 12 MELTFIQVDAFPADSPFGQNPAAVCPDLAWLDERLQAIABENNLSETPAV-VGRDGD--Y 68
QY 61 HATHTRTETIPFAGHPVTVGASWMLRR--GTPINTLQVPAIGIVQSYHGDLPATISARS 117
DB 69 RLRWTFQVEVDLCGHATLATAWLIHKLDASVPLAFATRSGLSTRREDSLAMPFPA 128
QY 118 EWAPEFAIHDLSDIALAADPADPDDIAHYLTWTDR----- 156
DB 129 K-RPEPCATPDGLLEALGIAEAVLKTDDYLWVVDDEKTIALLADPARKGLPKRGV 185
QY 157 --SAGSLR---AAMPAAANGVTDEATGAATITDYLSDLTIT-----QG---KGS 202
DB 186 AVTARSORFDFVSWFSPNNGVNDPVTGASHTSLAPYMAQRKLTSLAEOGGARGRL 245
QY 203 IHTWSPGQWVRVAGRVVSDGVAQL 227
DB 246 -----ECVYR-GERVVISGKAL 262

RESULT 15
AB078267
ID AB078267 standard; protein; 335 AA.
XX
AC AB078267;
XX
DT 29-JUL-2004 (first entry)

XX
DE Pseudomonas aeruginosa polypeptide #10442.
XX
KW Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.
XX
OS Pseudomonas aeruginosa.
XX
PN US6551795-B1.
XX
PD 22-APR-2003.
XX
PF 18-FEB-1999; 99US-00252991.
XX
PR 18-FEB-1998; 98US-0074788P.
XX 27-JUL-1998; 98US-0094190P.
XX
PA (GENO-) GENOME THERAPEUTICS CORP.
PI Rubenfield MJ, Nolling J, Deloughery C, Bush D;
XX WPI; 2003-615309/58.
DR N-PSDB; ABD11838.
XX
XX
PT Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
XX useful as molecular targets for diagnostics, prophylaxis and treatment of
XX pathological conditions resulting from bacterial infection.
XX
PS Disclosure; SEQ ID NO 27013; 455bp; English.
XX
CC The invention relates to Pseudomonas aeruginosa polypeptides and the
CC polynucleotides encoding them. The sequences are useful in diagnosis and
CC therapy of pathological conditions, as molecular targets for diagnostics,
CC prophylaxis and treatment of pathological conditions resulting from a
CC bacterial infection, for evaluating a compound, such as a polypeptide,
CC for the ability to bind a P. aeruginosa nucleic acid, as components of
CC effective antibacterial targets, as targets for antibacterial drugs,
CC including anti-P. aeruginosa drugs, as templates for recombinant
CC production of P. aeruginosa-derived peptides or polypeptides, as target
CC components for diagnosis and/or treatment of P. aeruginosa-caused
CC infection, and in detection of P. aeruginosa sequences or other sequences
CC of Pseudomonas species using biochip technology. Sequences AB067826-
CC AB084396 represent P. aeruginosa polypeptides of the invention. Note: The
CC sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format from USPTO at
CC seqdata.uspto.gov/sequence.html
XX
XX
SQ Sequence 335 AA;
Query Match 9.9%; Score 118; DB 7; Length 335;
Best Local Similarity 23.2%; Pred. No. 0.0033;
Matches 61; Conservative 34; Mismatches 98; Indels 70; Gaps 10;
QY 19 GNPLGVY-NASKVEHNRQQLAASGSEITFPVLDSPGSGTGAHTHTRTETIPFAGHP 77
DB 69 GNGLAVFDDASTLDLDAAMQMTRELRQFESIFL-LPDDDPRAVRARFTLEEBELPFAGHP 127
QY 78 TVGAS-----WTLRERGTPINTLQVPAIGIVQSYHGD-----TAIS 114
DB 128 LIGAAPALHLLRGQDNEQHTTLHASKSVLRSVRG--SGFAENDQGRAPFAGATPDA 184
QY 115 ARSEW-APEFAIHDLSDIALAADPADPDDIAHYLTWTDRSAGSLRA-----R 164
DB 185 GTGWRFAEAFSL-----SANDLSGHPRPVYSTGLPYLLPTABALGARGVNDLQALDK 240
QY 165 MFA-----NLGVTDEATGAATITDYLSDLTITTGKSLIY----- 204
DB 241 LGAAAFVTLIDVDEREGRTWDSLGLVEDVATGSAAGPAAVAVLYEGLAARGEPVYLHGRF 300
QY 205 -----TWSPEQWVRVAGRV 219
DB 301 LERPSRLDVQVATDGSVRVGHV 323

Search completed: April 14, 2006, 17:18:29
Job time : 91.341 secB

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: April 14, 2006, 17:19:08 ; Search time 13.3694 Seconds
(without alignments)
1640.866 Million cell updates/sec

Title: US-10-620-246-66

Perfect score: 1192
Sequence: 1 MAIEVSVLRVFTDSNFGN.....PEGWVRVAGRVSDVAQOLD 228

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1192	100.0	228	2 G70532	hypothetical prote
2	407	34.1	214	2 T50577	hypothetical prote
3	205.5	17.2	314	2 S75237	hypothetical prote
4	184.5	15.5	305	2 AB2742	hypothetical prote
5	180.5	15.1	333	2 A97523	hypothetical prote
6	161.5	13.5	276	2 AE3481	N-methyl-D-asparta
7	156	13.1	290	2 AF3234	phenazine biosynth
8	154.5	13.0	294	2 S46781	hypothetical prote
9	151.5	12.7	278	2 B83119	probable phenazine
10	148.5	12.5	308	2 AH3330	phenazine biosynth
11	145	12.2	301	2 H84192	hypothetical prote
12	144.5	12.1	281	2 AI3080	phenazine biosynth
13	144.5	12.1	290	2 H98205	hypothetical prote
14	131.5	11.0	290	2 G95273	conserved hypochet
15	131.5	11.0	295	2 F83693	hypothetical prote
16	130.5	10.9	293	2 C69800	conserved hypochet
17	126	10.6	296	2 S44972	lmbx protein - Str
18	125	10.5	259	2 AB3301	hypothetical prote
19	124	10.4	278	2 D83474	hypothetical prote
20	123	10.3	284	2 F83056	hypothetical prote
21	112.5	9.4	297	2 C64899	yddB protein - Bsc
22	112.5	9.4	297	2 C90887	hypothetical prote
23	112.5	9.4	302	2 F85730	hypothetical prote
24	112	9.4	302	2 F87323	probable epimerase
25	104.5	8.8	404	2 A75285	cinA protein - Del
26	104	8.7	279	2 F82463	conserved hypochet
27	104	8.7	282	2 AF1530	conserved hypochet
28	103	8.6	286	2 D86163	FKS3.19 protein -
29	101	8.5	316	2 T24413	hypothetical prote

30	100.5	8.4	308	2 E75409	probable antibiotic
31	99	8.3	284	2 F70024	diaminopimelate ep
32	97.5	8.2	282	2 AB1173	conserved hypochet
33	96.5	8.1	275	2 C87648	phenazine biosynth
34	95.5	8.0	776	2 T29064	hyaluronate lyase
35	95	8.0	251	2 E83199	conserved hypochet
36	95	8.0	273	2 AB2049	hypothetical prote
37	95	8.0	294	2 D85036	hypothetical prote
38	94.5	7.9	321	2 T24410	hypothetical prote
39	91.5	7.7	297	2 AF0671	conserved hypochet
40	91	7.6	728	2 AF3299	malate synthase (B
41	91	7.6	1145	2 S37136	structural polypro
42	91	7.6	15281	2 S41309	cyclosporin synthe
43	90.5	7.6	2591	2 T30288	distamycin I by
44	89	7.5	3413	2 T17467	rifamycin polyketi
45	88.5	7.4	572	2 JC7833	kumamolysin precur

ALIGNMENTS

RESULT 1

G70532
hypothetical protein RV2716 - Mycobacterium tuberculosis (strain H37RV)
N/Alternate names: thymidylate synthase (EC 2.1.1.45) [misidentification]
C/Species: Mycobacterium tuberculosis
C/Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C/Accession: G70532; S21832
R/Cole, S.T.; Brosch, R.; Parhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, I.; Connor, R.; Davies, R.; Devlin, K.; Felwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Ruter, S.; Seeger, K.; Skelton, S.; Squares, S. Nature 393, 537-544, 1998
A/Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A/Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A/Reference number: A70500; PMID:9825987; PMID:9634230
A/Accession: G70532
A/Status: nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-228 <COL>
A/Cross-references: UNIPROT:007215; UNIPARC:UP1000013BFB5; GB:Z96072; GB:AL123456; NID:
A/Experimental source: strain H37RV
R/Patki, A.H.; Dale, J.W.
Submitted to the EMBL Data Library, April 1991
A/Reference number: S21832
A/Accession: S21832
A/Molecule type: DNA
A/Residues: MSAGVTQDVNIVRSLSPKSGRAMLR, 'LPVLPAVQAGRIVGQGRHNGARPAQ', 'CAHNGRP,
RCASPGMR', 'TTVTGSADTRVAVALAIGSNERSATLITG', 'SVRALPQVVGSHG', '3-228 <PAT>
A/Cross-references: UNIPARC:UP1000016FBC6; EMBL:X59273; NID:944681; PIDN:CAA41963.1; PI
C/Genetics:
A/Gene: RV2716

Query Match	100.0%	Score 1192;	DB 2;	Length 228;
Best Local Similarity	100.0%	Pred. No. 2.1e-97;		
Matches 228;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MAIEVSVLRVFTDSNFGNPLGVINASKVHEHDDROOLAAQSGXSEITFYDLPSPGSTTA	60	
DB	1	MAIEVSVLRVFTDSNFGNPLGVINASKVHEHDDROOLAAQSGXSEITFYDLPSPGSTTA	60	
QY	61	HATHTPRTETPAGHPYVQASWMLERKGPINTLOVPAGIVQSVYHGDTTASRSEWA	120	
DB	61	HATHTPRTETPAGHPYVQASWMLERKGPINTLOVPAGIVQSVYHGDTTASRSEWA	120	
QY	121	PERAHDLSLDAALADPPDDTAHTLMTTDSAGSLRARMRAANGTDEDEATGA	180	
DB	121	PERAHDLSLDAALADPPDDTAHTLMTTDSAGSLRARMRAANGTDEDEATGA	180	
QY	181	AAIRITDYSRDLITIQGKSLHTTWSPEGVAVAGRVSDVAQOLD	228	
DB	181	AAIRITDYSRDLITIQGKSLHTTWSPEGVAVAGRVSDVAQOLD	228	

A:Cross-references: SGD:S0001071
A:Map position: 8R

Query Match 13.0%; Score 154.5; DB 2; Length 294;
Best Local Similarity 24.8%; Pred. No. 4.6e-06;
Matches 61; Conservative 38; Mismatches 88; Indels 59; Gaps 10;

QY 10 VFTSDGNFGNPGVGIN-----ASKVEHNRDQQLAASGSETIPVDLSPGSTTAHAT 64
DB 12 VFTSEK-FMGNPVAVFNFLEIDENRVSQBELQAIAMWTNSETPLFKPESDKYDKLRI 70
QY 65 HRPRTIPRAGHTTGAASWMLR--RGTPINTL--QVPAIGVSVHGDLTALSA----- 115
DB 71 FPRSELPRAGHPTIGSCAFLEPTKNTATSTLVQECKIGAVPITINEGLISFKAPMADY 130
QY 116 ---RSEW-----APEFAIHDLSDLAADPADP-----P 143
DB 131 ESISSEMIDYKALIGKPKPPALHTGEMVALVEDMFCFPMNP-NFAMLAHQTK 189
QY 144 DDIAHYLMTWTRDSA---GSLRAMPAMNGVTEDEATGAARITDYL-----SRD 192
DB 190 NDHVGIIAGPKKEAIAIKNSYEMRAFAVPINVVEDPVCSSVALARYLQEVYKFEKTD 249
QY 193 LTTTGG 198
DB 250 ITTSEG 255

RESULT 9

B83119 probable phenazine biosynthesis protein PA4215 [imported] - Pseudomonas aeruginosa (strain C:\Species: Pseudomonas aeruginosa
C:\Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C:\Accession: B83119; B83408
C:\Author: C.K.; Pham, X.O.; Ewlin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Bradman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim, J.; Loty, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:\Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathogen
A:\Reference number: A82950; PMID:20437337; PMID:10984043
A:\Accession: B83119
A:\Status: preliminary
A:\Molecule type: DNA
A:\Residues: 1-278 <STO>
A:\Cross-references: UNIPROT:O69754; UNIPARC:UPI0000131A7B; GB:AE004838; GB:AE004091; NID:A:\Experimental source: strain PA01
A:\Accession: B83408
A:\Status: preliminary
A:\Molecule type: DNA
A:\Residues: 1-278 <STO>
A:\Cross-references: UNIPARC:UPI0000131A7B; GB:AE004616; GB:AE004091; NID:G9947890; PIDN:A:\Experimental source: strain PA01
C:\Genetics:
A:\Gene: PA4215; PA1904

Query Match 12.7%; Score 151.5; DB 2; Length 278;
Best Local Similarity 28.1%; Pred. No. 7.8e-06;
Matches 62; Conservative 26; Mismatches 88; Indels 45; Gaps 9;

QY 19 GNPPLG-INAASKENHNRDQQLAASGSETIPVDLSPGSTTAHATHTPTTPRAGH 77
DB 17 GNPVAFVPCDDISGGRMOMAREMNISESTFVLRRPODG-DARIRLFPVNLPRAGH 75
QY 78 TVGASWML-RERGTPTNTLQVPAIGVSVH---GDLTALSARS---EW----- 119
DB 76 LLSATALAGETDKRLFLETRMGTVPFALERODGKVAACSMQOPLTFHESRPAELIA 135
QY 120 -----APEFAIHDLSDLAADPADP-----ADPDDIAHYLMTWTRDSA 158
DB 136 ALGLKSTPIIVYRNGRHRVPGLSVALSLHPDRALCPFP-DLANVCFAAGR-- 192
QY 159 GSIRAMPAMNGVTEDEATGAARITDYLSRDLTTTGG 199

DB 193 -HMRSMESPANGVEDAATGSAAGPLAHLARHROPYGO 232

RESULT 10

AH3330 phenazine biosynthesis protein phzF [imported] - Brucella melitensis (strain 16M)
C:\Species: Brucella melitensis
C:\Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 09-Jul-2004
C:\Accession: AH3330
R:\Belvecchio, V.G.; Kaputal, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Loeb, T.; Ivanova, J.; Mazur, M.; Goldstein, E.; Selkov, B.; Elzer, P.H.; Hagler, S.; O'Callaghan, D.; Leteser, Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A:\Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis
A:\Reference number: AD3252; PMID:11756688
A:\Accession: AH3330
A:\Status: preliminary
A:\Molecule type: DNA
A:\Residues: 1-308 <KUR>
A:\Cross-references: UNIPROT:O8Y115; UNIPARC:UPI0000057D60; GB:AE008917; PIDN:AAJ51811.1,
A:\Experimental source: strain 16M
C:\Genetics:
A:\Gene: BMEI0630
A:\Map position: I

Query Match 12.5%; Score 148.5; DB 2; Length 308;
Best Local Similarity 24.5%; Pred. No. 1.6e-05;
Matches 66; Conservative 31; Mismatches 95; Indels 77; Gaps 11;

QY 7 VLRVFTSDGNFGNPGVGI--NASKVEHNRDQQLAASGSETIPVDLSPGSTTAHAT 65
DB 13 VEDVFADK-ALAGNPLATVHDEGLTDACMQAIAREFNLSFTVFIAPNAPHEASVRIF 71
QY 66 TRTEIPRAGHTTGAASWMLRERGT-----INTLQVPAIGVSVHGDLTALS---- 114
DB 72 TRAVELPRAGHTTGAASWMLRERGT-----INTLQVPAIGVSVHGDLTALS---- 131
QY 115 -----ARSEWAPPA-----IHDDSLDLA 137
DB 132 PRLEPQVDYKIRERBAALIGKTHEIGFENHPGVASGTPCLLVPHNL-----IAA 186
QY 138 ---DPADPDDIAH-----YLM-TWTRDASGLRAMPAMNGVTEDEATGAAR 184
DB 187 KVISIDPVYSESLPHVGSRLPIYVCRFTILFNSPHARMVIGANVEDPATGSAVAA 246
QY 185 ITDYLSRDLTTTGGKSLHTTWSPEGW 213
DB 247 FAGMIQNDKPYDG-----SSQW 266

RESULT 11

H84192 hypothetical protein Vng0332c [imported] - Halobacterium sp. NRC-1
C:\Species: Halobacterium sp. NRC-1
C:\Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C:\Accession: H84192
R:\Ng, W.V.; Kennedy, S.P.; Mahataas, G.G.; Bergquist, B.; Pan, M.; Shukla, H.D.; Laakey, J.; Leitbauer, B.; Keller, K.; Cruz, R.; Danon, M.J.; Hought, D.W.; Maddocks, D.G.; Jable, Jung, K.H.; Alam, M.; Freltas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A:\Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebdardt, H.; Lowe, T.M.; Li, A.; Cross-references: UNIPROT:Q9HS99; UNIPARC:UPI00000635E1; GB:AE004437; NID:G10579960; PIDN:VNG0332C
A:\Status: preliminary
A:\Molecule type: DNA
A:\Residues: 1-301 <STO>
A:\Cross-references: UNIPROT:Q9HS99; UNIPARC:UPI00000635E1; GB:AE004437; NID:G10579960; PIDN:VNG0332C
A:\Gene: VNG0332C

Query Match 12.2%; Score 145; DB 2; Length 301;
Best Local Similarity 26.8%; Pred. No. 3.2e-05;
Matches 77; Conservative 32; Mismatches 96; Indels 82; Gaps 17;

QY 11 FTSDGNGFNPLGVI-NASKEHEDRQOLAAGSGYSEITFVDLPSPGSTTAHATHTPT 69
 DB 12 FTD-DSCAGNAAGVVPDAGGLTDQMOAIAELGASSTAF--LRPSAADRRLRYFPPT 68
 QY 70 EIPFAGHPYTGASWMLRERGTPT--NTLOVPAGI-----VQV 104
 DB 69 EVDLCGHATTAHAHLVDGALAAAGHTLTETNGVLDIEVAEDGTWMAQNPTVREVEL 128
 QY 105 SYH-----GDLTAISARSEMNP-----EPAIHDLSDLAADPADF-----P 143
 DB 129 SYRVVATETEAALAGASDDIPLAVADTGLPFLIPIYTLSDLGADP--DFDAVEALA 187
 QY 144 DDIAH--YLMWT--TDRSAGSLRAMPANLGVTEDEATGAARITDYLSR----- 191
 DB 188 DVGAGVWVFSPDALDR-ASTAHGRAMVPGVGVDEDPVTGTAAGAAGATLARGAFDGD 246
 QY 192 ---DLITTOGKSLIHTTWSPEGWVRVAGRVSD-----GVAQLD 228
 DB 247 APDEMVEEQG-----HEVDRP-GRVRY--RVAGDAPWVGKGVSTLD 285

RESULT 12

AI3080
 phenazine biosynthesis protein [imported] - Agrobacterium tumefaciens (strain C58, Dupont C/Specties: Agrobacterium tumefaciens
 C/Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
 C/Accession: AI3080
 R/Wood, D.W.; Seubald, J.C.; Kaul, R.; Monke, D.; Chen, L.; Wood, G.B.; Chen, Y.; Moo, L. et al., G.; Giller, W.; Grant, C.; Guenther, D.; Kutyavlin, T.; Levy, R.; Li, M.; McCell, S. et al., P.; Romero, P.; Zhang, S.
 Science 294, 2317-2323, 2001
 A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, A.; Biddle, P.
 A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
 A:Reference number: AB2577; MUID:21608550; PMID:11743193
 A:Accession: AI3080
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-281 <KUR>
 A:Cross-references: UNIPROT:Q8U830; UNIPARC:UP100001648FC; GB:AE008689; PTDN:AA45063.1
 A:Experimental source: strain C58 (Dupont)
 C:Genetics:
 A:Gene: phzC
 A:Map position: linear chromosome

Query Match 12.1%; Score 144.5; DB 2; Length 281;
 Best Local Similarity 27.4%; Pred. No. 3.3e-05;
 Matches 71; Conservative 32; Mismatches 105; Indels 51; Gaps 13;
 QY 1 MAIEVSVLRVFTSDGNGFNPLGVI-NASKEHEDRQOLAAGSGYSEITFVDLPSPGSTT 59
 DB 1 MTIAIHMDIVF-GSGPLSGNPLAVVVGADKLTDEMQRLTRWNLSTTLLPPTDPRAD 59
 QY 60 AHATHTPTREIPFAGHPYTGASW-MLRERGTPTINTLOV-----PAGIVQVSYHGDLTALS 114
 DB 60 YRVRIISLDREMPFAGHPYTGASWMLRERGTPTINTLOV-----PAGIVQVSYHGDLTALS 109
 QY 115 ARSEMAPBEPAIHDLSDLAADPADFPDDI-----AHYL--WTWDRSAGSLRAMP 166
 DB 110 RRAQGRISFA--APLIRSGAPTPALBERALQLLGIEADVDVDAAMIDNGPGLGVRLA 166
 QY 167 AA-----NLGVTEDEATGA-AAIRITDYLSRDL-TTQGS--KGSILHTT 206
 DB 167 SAKKVLSTDPVRSMPGRIDVGVGPHRAGEAFAFVRAFPSDHGAIAEDPVTGSL--N 223
 QY 207 WSPGWRVAVGRVVSQV 225
 DB 224 ASLAQWLFATGQWVAADYVA 242

RESULT 13

H98205

hypothetical protein AGR_L_1187 [imported] - Agrobacterium tumefaciens (strain C58, Cer C/Specties: Agrobacterium tumefaciens
 C/Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 09-Jul-2004
 C/Accession: H98205
 R/Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Onorollo, B.; Goldman, A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lippae, C.; Markelz, B. et al., Science 294, 2323-2328, 2001
 A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tumefaciens
 A:Reference number: A97359; MUID:21608551; PMID:11743194
 A:Accession: H98205
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-290 <KUR>
 A:Cross-references: UNIPROT:Q8U830; UNIPARC:UP100000242DF; GB:AE007870; PTDN:AAK69170.1
 C:Genetics:
 A:Gene: AGR_L_1187
 A:Map position: linear chromosome

Query Match 12.1%; Score 144.5; DB 2; Length 290;
 Best Local Similarity 27.4%; Pred. No. 3.4e-05;
 Matches 71; Conservative 32; Mismatches 105; Indels 51; Gaps 13;

QY 1 MAIEVSVLRVFTSDGNGFNPLGVI-NASKEHEDRQOLAAGSGYSEITFVDLPSPGSTT 59
 DB 10 MTIAIHMDIVF-GSGPLSGNPLAVVVGADKLTDEMQRLTRWNLSTTLLPPTDPRAD 68
 QY 60 AHATHTPTREIPFAGHPYTGASW-MLRERGTPTINTLOV-----PAGIVQVSYHGDLTALS 114
 DB 69 YRVRIISLDREMPFAGHPYTGASWMLRERGTPTINTLOV-----PAGIVQVSYHGDLTALS 118
 QY 115 ARSEMAPBEPAIHDLSDLAADPADFPDDI-----AHYL--WTWDRSAGSLRAMP 166
 DB 119 RRAQGRISFA--APLIRSGAPTPALBERALQLLGIEADVDVDAAMIDNGPGLGVRLA 175
 QY 167 AA-----NLGVTEDEATGA-AAIRITDYLSRDL-TTQGS--KGSILHTT 206
 DB 176 SAKKVLSTDPVRSMPGRIDVGVGPHRAGEAFAFVRAFPSDHGAIAEDPVTGSL--N 232
 QY 207 WSPGWRVAVGRVVSQV 225
 DB 233 ASLAQWLFATGQWVAADYVA 251

RESULT 14

G95273
 conserved hypothetical protein Sma0180 [imported] - Sinorhizobium meliloti (strain 1021.
 C/Specties: Sinorhizobium meliloti
 C/Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 09-Jul-2004
 C/Accession: G95273

R/Barnett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; Bow, J.; Kalman, S.; Keating, D.H.; Palm, C.; Peck, M.C.; Surzycki, R.; Welle, D.H.; Yeh, K.C. et al., Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001
 A:Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium meliloti.
 A:Reference number: A95262; MUID:21396509; PMID:11483432
 A:Accession: G95273
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-290 <KUR>
 A:Cross-references: UNIPROT:Q930U9; UNIPARC:UP100000CA9F7; GB:AE006469; PTDN:AAK64753.1
 A:Experimental source: strain 1021, megaplasmid pSymA
 R/Gallbert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler, L.; Hyman, R.W.; Jones, T.; Davie, R.W.; Dreano, S.; Federpiet, N.A.; Fisher, R.F. et al., Science 293, 668-672, 2001
 A:Authors: Kalman, D.; Kalman, M.L.; Kalman, S.; Keating, D.H.; Klase, E.; Komp, C.; Lelaurie, P.; Vandenberg, M.; Vorholter, F.J.; Weidner, S.; Welle, D.H.; Wong, K.; Yen, K. et al., Science 293, 668-672, 2001
 A:Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
 A:Reference number: A96039; MUID:21368234; PMID:11474104
 A:Contents: annotation
 C:Genetics:
 A:Gene: Sma0180
 A:Genome: plasmid

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: April 14, 2006, 17:04:10 ; Search time 84.1245 Seconds
(without alignments)
1912.171 Million cell updates/sec

Title: US-10-620-246-66

Perfect score: 1192
Sequence: 1 MAIEVSVLRYFTSDGNFGN.....PEGWVRVAGRVSDEVAQOLD 228

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UnIProt_05.80:.*
1: uniprot_sprot:.*
2: uniprot_trembl:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1192	100.0	228	1 Y2716_MYCTU	Pa557 mycobacteri
2	1192	100.0	228	1 Y2735_MYCBO	Pa558 mycobacteri
3	559.5	46.9	215	2 Q5Y752_NOCFA	O5Y752 nocardia fa
4	427	35.8	214	2 Q82X66_STRAM	Q82X66 streptomyc
5	407	34.1	214	2 Q9RKR0_STRCO	Q9RKR0 streptomyc
6	212.5	17.4	302	2 Q6N501_RHOBA	Q6N501 rhodospseudo
7	207.5	17.4	305	2 Q92M77_RHIME	Q92M77 rhizobium m
8	205.5	17.2	306	1 Y1019_SYNY3	P71315 synechocyst
9	199	16.7	292	2 Q4UWY3_XANCP	Q4UWY3 xanthomonas
10	199	16.7	292	2 Q8P771_XANCP	Q8P771 xanthomonas
11	199	16.7	873	2 Q5AKH0_EMERT	Q5AKH0 aspergillus
12	198.5	16.7	292	2 Q5UZA6_HALNA	Q5UZA6 halocaula
13	196	16.4	286	2 Q4LTD6_9BURK	Q4LTD6 burkholderi
14	196	16.4	309	2 Q4I2T5_GIBZE	Q4I2T5 gibberella
15	194	16.3	292	2 Q8P115_XANAC	Q8P115 xanthomonas
16	191.5	16.1	279	2 Q5N116_ZYMMO	Q5N116 zymomonas m
17	186	15.6	278	2 Q7NZL1_CHRYO	Q7NZL1 chromobacte
18	185.5	15.6	289	1 Y1503_PASMO	Q9CKJ2 pasteurella
19	185.5	15.6	289	1 Q9RJO0_STRCO	Q9RJO0 streptomyc
20	185.5	15.6	299	2 Q9HHG6_HALSA	Q9HHG6 halobacteri
21	185	15.5	313	2 Q4HNN2_GIBZE	Q4HNN2 gibberella
22	184.5	15.5	305	2 Q8UFO1_AGRIS	Q8UFO1 agrobacteri
23	180.5	15.1	333	2 Q7C264_AGRIS	Q7C264 agrobacteri
24	176.5	14.8	295	2 Q5NL41_ZYMMO	Q5NL41 zymomonas m
25	175.5	14.7	329	2 Q5H2Y5_XANOR	Q5H2Y5 xanthomonas
26	173.5	14.6	330	2 Q4W9M8_ASPTU	Q4W9M8 aspergillus
27	170.5	14.3	293	2 Q7WSU9_BORPA	Q7WSU9 bordetella
28	170.5	14.3	293	2 Q7W6K0_BORPA	Q7W6K0 bordetella
29	170.5	14.3	293	2 Q7VYZ3_BORBS	Q7VYZ3 bordetella
30	170	14.3	305	2 Q92NM9_RHIME	Q92NM9 rhizobium m
31	168	14.1	309	2 Q891J1_BRABA	Q891J1 bradyrhizob

32	167	14.0	296	1 A8S1_SCHPO	Q8H113 echinosacch
33	165.5	13.9	360	2 Q4NMA5_9MIC	Q4NMA5 arthrobacte
34	161.5	13.5	276	2 Q8G339_BRUCU	Q8G339 brucella su
35	161.5	13.5	276	2 Q8YEP3_BRUMS	Q8YEP3 brucella me
36	160.5	13.5	291	2 Q7WAP5_BORPA	Q7WAP5 bordetella
37	160.5	13.5	291	2 Q7WU9_BORBR	Q7WU9 bordetella
38	159.5	13.4	278	2 Q5LKG2_SILPO	Q5LKG2 silicibacte
39	159.5	13.4	280	2 Q8NM35_CORGL	Q8NM35 corynebacte
40	157.5	13.2	298	2 Q6D274_BRWCT	Q6D274 erwinia car
41	156.5	13.1	276	2 Q57F08_BRUBB	Q57F08 brucella ab
42	156	13.1	290	2 Q8U679_AGRIS	Q8U679 agrobacteri
43	154.5	13.0	285	2 Q8Y2T2_RALSO	Q8Y2T2 ralscristia
44	154.5	13.0	294	1 YH19_YEAST	YH19 yeast
45	154.5	13.0	297	2 Q6BVN8_DEBRA	Q6BVN8 debaryomyce

ALIGNMENTS

RESULT 1
Y2716_MYCTU STANDARD; PRT; 228 AA.
ID Y2716_MYCTU
AC PA557; 007215; P28176;
DT 01-JUL-1993 (Rel. 26, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Hypothetical protein RV2716/MT2789.
GN OrderedLocName=RV2716; MT2789; ORFName=MYC05A6.37;
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium;
OC Mycobacterium tuberculosis complex.
OX NCBI_TaxID=1773;
[1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=H37Rv;
RX MEDLINE=98295987; PubMed=9634230; DOI=10.1038/31159;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C.M.,
RA Harris D.E., Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III,
RA Tekala F., Badcock K., Basham D., Brown D., Chillingworth T.,
RA Connor R., Davies R.M., Devlin K., Felwell T., Gentles S., Hamlin N.,
RA Holroyd S., Hornsby T., Jagels K., Krogh A., McLean J., Moule S.,
RA Murphy L.D., Oliver S., Osborne J., Quail M.A., Rajandream M.A.,
RA Rogers J.D., Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sultison J.E., Taylor K., Whitehead S., Barrett B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence.";
RL Nature 393:537-544(1998).
[2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=CDC 1551 / Oshkosh;
RX MEDLINE=22206494; PubMed=12218036;
RA DOI=10.1128/JB.184.19.5479-5490.2002;
RA Fleischnann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J.D., Debey R.T., Dodson R.J., Gaitan M.L., Haft D.H.,
RA Hickey B.K., Kolonay J.F., Nelson W.C., Unanue L.A., Ermolaeva M.D.,
RA Salzberg S.L., Delcher A., Uitterback T.R., Weidman J.F., Khouri H.M.,
RA Gill J., Mikula A., Bishai W., Jacobs W.R. Jr., Venter J.C.,
RA Fraser C.M.;
RT "Whole-genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains.";
RL J. Bacteriol. 184:5479-5490(2002).
[3]
RP PRELIMINARY NUCLEOTIDE SEQUENCE.
RC STRAIN=Isolate 50410;
RA Patki A.H., Dale J.W.;
RL Submitted (APR-1991) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: Ref.3 sequence was incorrect in two aspects; its sequence
CC was incorrect due to several framehifts (including one that fuses
CC RV2715 and RV2716 into one putative protein) and it was wrongly
CC assigned to be a thymidylate synthase.
CC This Swiss-Prot entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.

CC EMBL; BX842580; CAB09457.1; -; Genomic DNA.
CC EMBL; AE000516; AAK47105.1; -; Genomic DNA.
CC EMBL; X59273; CAA41963.1; ALT_FRAME; Genomic DNA.
CC PIR; G70532; G70532.
CC TIGR; W2769; -.

DR Tuberculin; Rv2716; -
DR Interpro; IPR003719; PhzC_PhzF.
DR PANTHER; PTHR13774; PhzC_PhzF; 1.
DR Pfam; PF02567; PhzC_PhzF; 1.
KW Complete proteome; Hypothetical protein.

SQ SEQUENCE 228 AA; 24574 MW; FEB2FD0DAB2A09B CRC64;

Query Match 100.0%; Score 1192; DB 1; Length 228;
Best Local Similarity 100.0%; Pred. No. 1,4e-93;
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAIEVSLRVFTSDGNFNGPLGVINASKVEHRDROQLAAGSGSETIFVLDLPSGSTTA 60
DB 1 MAIEVSLRVFTSDGNFNGPLGVINASKVEHRDROQLAAGSGSETIFVLDLPSGSTTA 60

QY 61 HATHTPRTETPPAGHPPTVGASWMLRERGTPIINTLQVPAGIVOVSYHGLTALSARSEWA 120
DB 61 HATHTPRTETPPAGHPPTVGASWMLRERGTPIINTLQVPAGIVOVSYHGLTALSARSEWA 120

QY 121 PEFAIHDLDSLDAALADPADPPDDIAHYLTMTWTDRSAGSLRARMFANLGVTEDEATGA 180
DB 121 PEFAIHDLDSLDAALADPADPPDDIAHYLTMTWTDRSAGSLRARMFANLGVTEDEATGA 180

QY 181 AAIRITDYSRLDTITQKGSLLHTTWSPEGWVAVAGRVSDGVAQLD 228
DB 181 AAIRITDYSRLDTITQKGSLLHTTWSPEGWVAVAGRVSDGVAQLD 228

RESULT 2

Y2735 MYCBO STANDARD; PRT; 228 AA.
ID T2735 MYCBO STANDARD; PRT; 228 AA.
AC F0A5G8; O07215; P28176;

DT 01-JUL-1993 (Rel. 26, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Hypothetical protein Mb2735.

GN OrderedLocuNames=Mb2735;

OS Mycobacterium bovis.

OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;

OC Corynebacterineae; Mycobacteriaceae; Mycobacterium;

OC Mycobacterium tuberculosis complex.

NCBI_TaxId=1765;

NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].

RC STRAIN=AF2122/97; PubMed=12788972; DOI=10.1073/pnas.1130426100;

RA Gardier T., Bidmeier K., Camus J.-C., Medina N., Mansoor H.,

RA Pryor M., Duthoy S., Grondin S., Lacroix C., Monsemp C., Simon S.,

RA Harris B., Atkin R., Doggett J., Mayer R., Keating L., Wheeler P.R.,

RA Parkhill J., Barrell B.G., Cole S.T., Gordon S.V., Hewison R.G.;

RT "The complete genome sequence of Mycobacterium bovis."

Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882(2003).

CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.

CC EMBL; BX248343; CAD94920.1; -; Genomic DNA.

DR Interpro; IPR003719; PhzC_PhzF.

DR PANTHER; PTHR13774; PhzC_PhzF; 1.

DR Pfam; PF02567; PhzC_PhzF; 1.

KW Complete proteome; Hypothetical protein.

SQ SEQUENCE 228 AA; 24574 MW; FEB2FD0DAB2A09B CRC64;

Query Match 100.0%; Score 1192; DB 1; Length 228;
Best Local Similarity 100.0%; Pred. No. 1,4e-93;
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAIEVSLRVFTSDGNFNGPLGVINASKVEHRDROQLAAGSGSETIFVLDLPSGSTTA 60
DB 1 MAIEVSLRVFTSDGNFNGPLGVINASKVEHRDROQLAAGSGSETIFVLDLPSGSTTA 60

QY 61 HATHTPRTETPPAGHPPTVGASWMLRERGTPIINTLQVPAGIVOVSYHGLTALSARSEWA 120
DB 61 HATHTPRTETPPAGHPPTVGASWMLRERGTPIINTLQVPAGIVOVSYHGLTALSARSEWA 120

QY 121 PEFAIHDLDSLDAALADPADPPDDIAHYLTMTWTDRSAGSLRARMFANLGVTEDEATGA 180
DB 121 PEFAIHDLDSLDAALADPADPPDDIAHYLTMTWTDRSAGSLRARMFANLGVTEDEATGA 180

QY 181 AAIRITDYSRLDTITQKGSLLHTTWSPEGWVAVAGRVSDGVAQLD 228
DB 181 AAIRITDYSRLDTITQKGSLLHTTWSPEGWVAVAGRVSDGVAQLD 228

RESULT 3

Q5Y752 NOCPA PRELIMINARY; PRT; 215 AA.
ID Q5Y752 NOCPA PRELIMINARY; PRT; 215 AA.
AC Q5Y752;

DT 25-OCT-2004 (TrEMBLrel. 28, Created)

DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)

DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)

DE Hypothetical protein.

GN OrderedLocuNames=nfa37910;

OS Nocardia farcinica.

OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;

OC Corynebacterineae; Nocardiaceae; Nocardia.

NCBI_TaxId=37329;

NUCLEOTIDE SEQUENCE.

RC STRAIN=IFM 10152;

RX PubMed=15466710; DOI=10.1073/pnas.0406410101;

RA Ishikawa J., Yamashita A., Mikami Y., Hoshino Y., Kurita H., Hotta K.,

RA Shiba T., Hattori M.;

RT "The complete genomic sequence of Nocardia farcinica IFM 10152."

Proc. Natl. Acad. Sci. U.S.A. 101:14925-14930(2004).

EMBL; AF006618; BAD58639.1; -; Genomic DNA.

DR GO; GO:0003824; F: catalytic activity; IEA.

DR GO; GO:0009058; P: biosynthesis; IEA.

DR Interpro; IPR003719; PhzC_PhzF.

DR PANTHER; PTHR13774; PhzC_PhzF; 1.

DR Pfam; PF02567; PhzC_PhzF; 1.

KW Complete proteome; Hypothetical protein.

SQ SEQUENCE 215 AA; 23024 MW; 4452A54B1C84D0B CRC64;

Query Match 46.9%; Score 559.5; DB 2; Length 215;
Best Local Similarity 48.4%; Pred. No. 1,4e-39;
Matches 109; Conservative 31; Mismatches 68; Indels 17; Gaps 4;

QY 10 VFTDSDGNFNGPLGVINASKVEHRDROQLAAGSGSETIFVLDLPSGSTTAHATITPRT 69
DB 1 MFVDSGRLGNALGIRAGADVADADQALNARGFSEYVV--ADVAVAYAKRITTPAM 58

QY 70 ELPFAGHPPTVGASWMLRERGTPIINTLQVPAGIVOVSYHGLTALSARSEWAPFAIHDL 129
DB 59 ELPFAGHPPTVGASWMLRERGTPIINTLQVPAGIVOVSYHGLTALSARSEWAPFAIHDL 118

QY 130 SLDALAALADPADP---PPDIA---HYLTMTWTDRSAGSLRARMFANLGVTEDEATGA 182
DB 119 -----DPABDELRLPDDFTGSPHYLMAMDEKMGSRSRFAFTMGIAEDAAATGA 170

QY 183 IRTDYLSDRLTITQKGSLLHTTWSPEGWVAVAGRVSDGVAQLD 227
DB 171 VALTALLRQGLKITQAGSOLFTEWSDGKVRIGRVVADSVLEI 215

RESULT 4

082AX6 STRAM

ID 082AX6 STRAM PRELIMINARY; PRT; 214 AA.

AC 082AX6;

DT 01-JUN-2003 (TReMBLrel. 24, Created)

DT 01-JUN-2003 (TReMBLrel. 24, Last sequence update)

DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)

DE Hypothetical protein.

GN OrderedLocusNames=SAV5929;

OS Streptomyces avermitilis.

OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;

OC Streptomycineae; Streptomycetaceae; Streptomyces.

NCBI_TaxID=33903;

[1] NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].

RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;

RX MEDLINE=21477403; PubMed=11572948; DOI=10.1073/pnas.211433198;

RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,

RA Shinoe M., Takahashi Y., Horikawa H., Nakazawa H., Osomoe T.,

RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.,

RT "Genome sequence of an industrial microorganism Streptomyces

RT avermitilis: deducing the ability of producing secondary

RT metabolites."

RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).

[2] NUCLEOTIDE SEQUENCE.

RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;

RX MEDLINE=22608306; PubMed=12692562; DOI=10.1038/nbt820;

RA Ikeda H., Ishikawa J., Hanamoto A., Shinoe M., Kikuchi H., Shiba T.,

RA Sakaki Y., Hattori M., Omura S.,

RT "Complete genome sequence and comparative analysis of the industrial

RT microorganism Streptomyces avermitilis."

RL Nat. Biotechnol. 21:526-531(2003).

DR EMBL; BA000030; BAC73641.1; -; Genomic DNA.

DR GO; GO:0003824; P:catalytic activity; IEA.

DR GO; GO:0009058; P:biosynthesis; IEA.

DR InterPro; IPR003719; PhzC PhzF.

DR PANTHER; PTHR13774; PhzC_PhzF; 1.

DR Pfam; PF02567; PhzC_PhzF; 1.

KW Complete proteome; Hypothetical protein.

SQ SEQUENCE 214 AA; 23077 MW; 0ED571AD9BPD7D98 CRC64;

Query Match 35.8%; Score 427; DB 2; Length 214;

Best Local Similarity 43.9%; Pred. No. 2.8e-26; Matches 98; Conservative 26; Mismatches 79; Indels 20; Gaps 6;

QY 4 EVSVLRVFTDSDGNFGNPLGVINASKY--EHRDRQOLAAGSYSETIFVDLPSPGSTTAH 61

DB 3 DYVNLVFCGPGGNGELGVNDSVLPFCADRGALAAKLGSEYTFVDDPERGV---- 58

QY 62 ATHTPRTETPFAGHPFTVGSWMLRERGTPINTLQVPAGIVQVSYHGDLTALISARSWAP 121

DB 59 IDYTTPLRLPFAGHPFCVGTAMLL-----DVPELVTRAGVVGTRLDSEFSGMIERPEMAP 113

QY 122 EFAIHLDLS---IDALAAADPADPPDDIAHYLTMTWDRSAGSLRARMF--AANGVTDEDA 177

DB 114 PRTLRQYAAVAAYVDLLVPPKGSV-----YAAWMBDEAAGRYARAFPGRDGIDEDDA 168

QY 178 TGAAAAIRITDYLRSRLITITGKSLIHTTWSPEGWVAVAGRV 220

DB 169 TGAALLLTGRLGRALNITIGRSGQILTAPQPGWVGVRVLT 211

RESULT 5

Q9KRK0 STRCO

ID Q9KRK0 STRCO PRELIMINARY; PRT; 214 AA.

AC Q9KRK0;

DT 01-MAY-2000 (TReMBLrel. 13, Created)

DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)

DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)

DE Hypothetical protein SCO2268.

GN OrderedLocusNames=SCO2268; ORFNames=SCC75A.14c;

OS Streptomyces coelicolor.

OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;

OC Streptomycineae; Streptomycetaceae; Streptomyces.

NCBI_TaxID=1902;

[1] NUCLEOTIDE SEQUENCE.

RC STRAIN=A3(2) / M145;

RX MEDLINE=21996410; PubMed=12000953; DOI=10.1038/417141a;

RA Bentley S.D., Chater K.F., Cerdano-Tarraga A.-M., Challis G.L.,

RA Thomson N.R., James K.D., Harris D.B., Quail M.A., Kleiser H.,

RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,

RA Cronin A., Frazer A., Goble A., Hidalgo J., Hornsby T., Howarth S.,

RA Huang C.-H., Kleiser T., Larke L., Murphy L.D., Oliver K., O'Neill S.,

RA Rabinowitsch R., Rajandream M.A., Rutherford K.M., Rutter S.,

RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,

RA Warren T., Wietzorrek A., Woodward J.R., Barrell B.G., Parkhill J.,

RA Hopwood D.A.;

RT "Complete genome sequence of the model actinomycete Streptomyces

RT coelicolor A3(2)."

RL Nature 417:141-147(2002).

DR EMBL; AL939112; CAB61714.1; -; Genomic DNA.

DR PIR; T50577; T50577.

DR GO; GO:0003824; P:catalytic activity; IEA.

DR GO; GO:0009058; P:biosynthesis; IEA.

DR InterPro; IPR003719; PhzC PhzF.

DR PANTHER; PTHR13774; PhzC_PhzF; 1.

DR Pfam; PF02567; PhzC_PhzF; 1.

KW Complete proteome; Hypothetical protein.

SQ SEQUENCE 214 AA; 23321 MW; A5A53BF5E9B9C07 CRC64;

Query Match 34.1%; Score 407; DB 2; Length 214;

Best Local Similarity 42.8%; Pred. No. 1.4e-26; Matches 95; Conservative 27; Mismatches 80; Indels 20; Gaps 6;

QY 4 EVSVLRVFTDSDGNFGNPLGVINASKY--EHRDRQOLAAGSYSETIFVDLPSPGSTTAH 61

DB 3 DYVNLVFCAPNGGNGELGVNDSVLPFPERRQELAAKLGSEYTFVDDPERGV---- 58

QY 62 ATHTPRTETPFAGHPFTVGSWMLRERGTPINTLQVPAGIVQVSYHGDLTALISARSWAP 121

DB 59 VDIYTTPLRLPFAGHPFCVGTAMLL-----DVPELVTRAGVVGARLDSEFSGMIERPEMAP 113

QY 122 EFAIHLDLS---IDSDALAAADPADPPDDIAHYLTMTWDRSAGSLRARMF--AANGVTDEDA 177

DB 114 PRTLRQYAAVAAYVDLLVPPKGSV-----YAAWMBDEAAGRYARAFPGRDGIDEDDA 168

QY 178 TGAAAAIRITDYLRSRLITITGKSLIHTTWSPEGWVAVAGRV 219

DB 169 TGAALLLTGRLGRALNITIGVGSQILTAPQPGWVGVRVLT 210

RESULT 6

Q6N501 RHOPA

ID Q6N501 RHOPA PRELIMINARY; PRT; 302 AA.

AC Q6N501;

DT 05-JUL-2004 (TReMBLrel. 27, Created)

DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)

DE Peniazine biosynthesis PhzC/PhzF protein.

GN Name=phzF; OrderedLocusNames=RP3182;

OS Rhodopseudomonas palustris.

OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;

OC Bradyrhizobiaceae; Rhodopseudomonas.

NCBI_TaxID=1076;

[1] NUCLEOTIDE SEQUENCE.

RC STRAIN=CGA009 / ATCC BAA-98;

RX PubMed=14704707; DOI=10.1038/nbt923;

RA Larimer F.W., Chain P., Hauber L., Lamerdin J.E., Malfatti S., Do L.,

RA Land M.L., Pelletier D.A., Beatty J.T., Lang A.S., Tabita F.R.,

RA Gibson J.L., Hanson T.E., Bobst C., Torres Y. Torres J.L., Peres C.,

RA Harrison F.H., Gibson J., Harwood C.S.;


```

QY 67 PRTEIPAGHPPTVGASW-WLRERGTINT-----LOVPAGIVQSVH---GDL----- 110
DB 67 PKRBLDAGHPPTGTYLLGLLOPPSPLVTTWQLESPVGLVPTLHYKQGLVQTELT 126
QY 111 -----TAISARS-----EMAPB-----FAIHDLSDIAL- AAAD 138
DB 127 AOLPOTDSAPSCEDLALLGLSLDQAGGHEYRQASGLPFIPLINEBALNRISFN 186
QY 139 PADPDDIAH-----YLTWTRDSAG-----SIRAMPANIGVTEDEATGAARITD 187
DB 187 PSWQNLACQWMDVCYCLAPDPSLGLSDNKLHGMFAPGIGIADPATGSGVALGG 246
QY 188 YL-----SRDLTTQSGK-----SLHTTMSRGG-----YVAGR--VVSQVADL 227
DB 247 YLDDRLDTFGSHHWQLEGGKALGRPSQLQTLVWQGGGIRAVKVAGRSVLSBGFNML 304

```

RESULT 9

```

Q4UWY3_XANCP PRELIMINARY; PRT; 292 AA.
ID Q4UWY3_XANCP PRELIMINARY; PRT; 292 AA.
AC Q4UWY3;
DT 13-SBP-2005 (Tremblrel. 31, Created)
DT 13-SBP-2005 (Tremblrel. 31, Last sequence update)
DE 13-SBP-2005 (Tremblrel. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNames=XC_1371;
OS Xanthomonas campestris pv. campestris str. 8004.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxId=314565;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=8004;
RA Qian W., Jia Y.-T., Ren S.-X., He Y.-Q., Feng J.-X., Lu L.-F.,
RA Sun Q.-H., Ying G., Tang D.-J., Wu W., Wang L.-F., Jiang B.-L.,
RA Zeng S.-Y., Gu W.-Y., Lu G., Kong L., Tian Y.-C., Yao Z.-J., Fu G.,
RA Chen B.-S., Fang R.-X., Qiang B.-Q., Chen Z., Zhao G.-P., Tang J.-L.,
RA He C.-Z.;
RA Submitted (FEB-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; CP000050; AAY48440.1; -; Genomic DNA.
KW Hypothetical protein.
SQ SEQUENCE 292 AA; 30927 MW; DB83DBBBE62636CA CRC64;

```

```

Query Match 16.7%; Score 199; DB 2; Length 292;
Best Local Similarity 29.0%; Pred. No. 1.2e-08;
Matches 82; Conservative 35; Mismatches 82; Indels 84; Gaps 15;
QY 8 LRVFTSDGNGFPLG-VINASKYEHDRQQLAAGSGYSETI-FVLLPSGSTTAATIH 65
DB 9 LDVFSNRPG-AGNPPLGVVDAGALSQOMQOIAAMNLSTVFPLPVASAGA-DYHIRIP 66
QY 66 TPRTIIPFAGHPVTGASWMLRERG---TPINTL--QVPAGIVQV----- 104
DB 67 TPRAELPFAGHPEVGAAWAKRTGLVGYRQDNRLYQCCAGVLPVDFRHDVVLVRLA 126
QY 105 -----SYH-----GDLTAISARSEMAPB-----AIHDLDSL 131
DB 127 PRAQSDGTGTHHAALAQVTRAFAMSAQAPALMNNNGPSWMLLELADEAAVRTAVPDLAI 186
QY 132 DALAADPADPFPDIDIAHYLTWTDRSAGSLRAMPFANIGVTEDEATGA---AIRITD 187
DB 187 AALTAASGA-----VGLATVAAAASDAEDLVVAFPCGCGIIPDPVPGSANACTAARL-- 239
QY 188 YLSRD-LITTOGKSLIHTTWSPEGWVAVAGRVVS-DGVAQLD 228
DB 240 -LAEDRLPGRGR-----YVAGQREVGRGVRVE 269

```

RESULT 10

```

Q8P771_XANCP PRELIMINARY; PRT; 292 AA.
ID Q8P771_XANCP PRELIMINARY; PRT; 292 AA.
AC Q8P771;
DT 01-OCT-2002 (Tremblrel. 22, Created)

```

```

DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE Hypothetical protein XCC2742.
GN OrderedLocustNames=XCC2742;
OS Xanthomonas campestris (pv. campestris).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxId=340;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 33913 / NCPPB 528;
RX MEDLINE=22022145; PubMed=12024217; DOI=10.1038/417459a;
RA da Silva A.C.R., Ferro J.A., Reinach P.C., Farah C.S., Furian L.R.,
RA Quaglini R.B., Monteiro-Vitorello C.B., Van Sluys M.A.,
RA Almeida N.F., Jr., Alves L.M.C., do Amaral A.M., Bertolini M.C.,
RA Camargo L.B.A., Camarotte G., Cannavan F., Cardozo J., Chambergo F.,
RA Clapiena L.P., Ciccarelli R.M.B., Coutinho L.L., Cursino-Santos J.R.,
RA El-Dorri H., Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Gruber A.,
RA Ferro M.I.T., Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA Martins B.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
RA Spindola L.A.F., Takita M.A., Tamura R.B., Teixeira B.C., Tezza R.I.D.,
RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
RA Secubal J.C., Kiteajima J.P.;
RT "Comparison of the genomes of two Xanthomonas pathogens with differing
RT host specificities."
RL Nature 417:459-463(2002).
DR EMBL; AB012387; AAM42014.1; -; Genomic DNA.
DR HSSP; P3757; ISDJ.
DR GO; GO:0003824; P:catalytic activity; IEA.
DR GO; GO:0009058; P:biosynthesis; IEA.
DR InterPro; IPR003719; PhzC_PhzF.
DR PANTHER; PTHR13774; PhzC_PhzF; 1.
DR Pfam; PF02567; PhzC_PhzF; 1.
DR TIGRPFAMs; TIGR00654; PhzF family; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 292 AA; 30927 MW; DB83DBBBE62636CA CRC64;

```

```

Query Match 16.7%; Score 199; DB 2; Length 292;
Best Local Similarity 29.0%; Pred. No. 1.2e-08;
Matches 82; Conservative 35; Mismatches 82; Indels 84; Gaps 15;
QY 8 LRVFTSDGNGFPLG-VINASKYEHDRQQLAAGSGYSETI-FVLLPSGSTTAATIH 65
DB 9 LDVFSNRPG-AGNPPLGVVDAGALSQOMQOIAAMNLSTVFPLPVASAGA-DYHIRIP 66
QY 66 TPRTIIPFAGHPVTGASWMLRERG---TPINTL--QVPAGIVQV----- 104
DB 67 TPRAELPFAGHPEVGAAWAKRTGLVGYRQDNRLYQCCAGVLPVDFRHDVVLVRLA 126
QY 105 -----SYH-----GDLTAISARSEMAPB-----AIHDLDSL 131
DB 127 PRAQSDGTGTHHAALAQVTRAFAMSAQAPALMNNNGPSWMLLELADEAAVRTAVPDLAI 186
QY 132 DALAADPADPFPDIDIAHYLTWTDRSAGSLRAMPFANIGVTEDEATGA---AIRITD 187
DB 187 AALTAASGA-----VGLATVAAAASDAEDLVVAFPCGCGIIPDPVPGSANACTAARL-- 239
QY 188 YLSRD-LITTOGKSLIHTTWSPEGWVAVAGRVVS-DGVAQLD 228
DB 240 -LAEDRLPGRGR-----YVAGQREVGRGVRVE 269

```

RESULT 11

```

Q5AXH0_EMENTI PRELIMINARY; PRT; 873 AA.
ID Q5AXH0_EMENTI PRELIMINARY; PRT; 873 AA.
AC Q5AXH0;
DT 10-MAY-2005 (Tremblrel. 30, Created)
DT 10-MAY-2005 (Tremblrel. 30, Last sequence update)
DT 10-MAY-2005 (Tremblrel. 30, Last annotation update)

```

Hypothetical protein.
 ORFNames=AN7010.2;
 OS Aspergillus nidulans FGSC A4.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 OC Eurotiiales; Trichocomaceae; Emericella.
 NCBI_TaxID=227321;
 OK NCB1_TaxID=227321;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=FGSC A4;
 RA Birren B., Nudman C., Abouellell A., Allen N., Anderson S.,
 Archach H.M., Barina N., Baetien V., Bloom T., Boguslavsky L.,
 Boukhalter B., Butler J., Calvo S.B., Camarata J., Chang J.,
 Choepl Y., Collimore A., Cook A., Cooke P., Corum B., Deatellano K.,
 Diaz J.S., Dodge S., Doolley K., Dorris L., Elkins T., Engels R.,
 Erickson J., Faro S., Ferreira P., Fitzgerald M., Gage D., Galagan J.,
 Gardyna S., Gierre S., Graham L., Grand-Pierre N., Hafez N.,
 Hagopian D., Hago B., Hall J., Horton L., Hulme W., Iliev I.,
 Jaffe D., Johnson R., Jones C., Kamal M., Katat A., Karatas A.,
 Kells C., Landers T., Levine R., Lindblad-Toh K., Liu G., Lui A.,
 Ma L.-J., Mabbitt R., Maclean C., Macdonald P., Major J., Manning J.,
 Mathews C., Muccelli E., McCarthy M., Meldrum J., Menus L.,
 Mithova T., Mlenga V., Murphy T., Naylor J., Nguyen C., Nicol R.,
 Nielsen C.B., Norbu C., O'Connor T., O'Donnell P., O'Neill D.,
 Oliver J., Peterson K., Phunhang P., Pierre N., Purcell S.,
 Rachupka A., Ramasamy U., Raymond C., Retta R., Rise C., Rogov P.,
 Roman J., Schauer S., Schupack R., Seaman S., Severy P., Smitrov S.,
 Smith C., Spencer B., Stange-Thomann N., Stojanovic N., Stubbs M.,
 Talamas J., Tesfaye S., Theodore J., Topham K., Travers M.,
 Vassiliev H., Venkataman V.S., Viel R., Vo A., Wang S., Wilson B.,
 Wu X., Wyman D., Young G., Zainoun J., Zembek L., Zimmer A., Zody M.,
 RA Lander E.;
 RA "Genome Sequence of Aspergillus nidulans";
 RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
 CC -!- CAUTION: The sequence shown here is derived from an
 EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 preliminary data.
 CC EMBL; AACD0100117; EAA61656.1; -!- Genomic DNA.
 DR Hypothetical protein.
 KW SEQUENCE 873 AA; 97928 MW; 38C15062ED48BC8 CRC64;

Query Match 16.7%; Score 199; DB 2; Length 873;
 Best Local Similarity 27.3%; Pred. No. 4.1e-08;
 Matches 78; Conservative 34; Mismatches 82; Indels 92; Gaps 12;

QY 3 IEVSVAVFPTDSGNNGNPGV-----NASKVHEHDDROOLAQSGYSETFV-DLPSGTTAAHTPTPTTIPFAGH 57
 DB 57 VDPVTLDVFTTKYE-GNPLAVVFLPAPSTQLTQROKOTIAKEFLSETIFVHSSGFS 115
 QY 58 TTAHATHTPTREIPRAPHPTVG-ASWMLRE--RGTPIPTLVPAIVGVSY-HGDLTAI 113
 DB 116 ESRKIDIFPGSEELPFAHGHTTGAASWPLHAPESKPVTRLWKADDFITFQNALGVV 175
 QY 114 SARSEWAPFAIH-----DLDSIDL 134
 DB 176 SARV--AAHVHHEHNOYPLDELRLYSPVYLTQSSIALISIVKMSQQLVELPSELV 233
 QY 135 AAADPADPDDIA-----HYLMTWDRSAGS--LRAPFAAMGVTEPB 176
 DB 234 GAVTTNNGSEKSSSYLDAGMAGMTTTFYVDVDELGRVITRTI--LGSLEDP 290
 QY 177 ATGAARITDYLSRDLITTOGKSLIHTTWSPEGWVAVGRVSD 222
 DB 291 ATGSAASGLTAVLS-----LKEGR-----AGRVVD 316

RESULT 12
 OSU2J6 HALMA PRELIMINARY; PRT; 292 AA.
 AC OSU2J6_7
 DT 01-FEB-2005 (TREMBlrel. 29, Created)
 DT 01-FEB-2005 (TREMBlrel. 29, Last sequence update)
 DT 01-FEB-2005 (TREMBlrel. 29, Last annotation update)
 DE Phenazine biosynthesis-like protein.

GN Name=phzC; OrderedLocustNames=trnAC2502;
 OS Haloarcula marismortui (Haloacterium marismortui);
 OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
 OC Halobacteriaceae; Haloarcula.
 NCBI_TaxID=2238;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=ATCC 43049;
 RX PubMed=15520287; DOI=10.1101/gr.2700304;
 RA Ballga N.S., Bonneau R., Facciotti M.T., Pan M., Guzman G.,
 RA Deutsch E.W., Shannon P., Chiu Y., Weng R.S., Gan R.R., Hung P.,
 RA Date S.V., Marcotte E., Hood L., Ng W.V.;
 RL "Genome sequence of Haloarcula marismortui: a halophilic archaeon from
 the Dead Sea.";
 RL Genome Res. 14:2221-2234(2004).
 DR EMBL; AY56297; AAU47307.1; -!- Genomic DNA.
 DR GO; GO:0003824; P: catalytic activity; IEA.
 DR GO; GO:0003056; P: biosynthesis; IEA.
 DR InterPro; IPR003719; PhzC_PhzF.
 DR InterPro; IPR001412; trna-lync-I.
 DR PANTHER; PTHR13774; PhzC_PhzF; 1.
 DR Pfam; PF02567; PhzC_PhzF; 1.
 DR TIGRfam; TIGR00654; PhzF family; 1.
 DR PROSITE; PS00178; AA_TRNA_LIGASE_1; UNKNOWN_1.
 KW Complete proteome.
 SQ SEQUENCE 292 AA; 31390 MW; 4A6BEE9915AFDA CRC64;

Query Match 16.7%; Score 198.5; DB 2; Length 292;
 Best Local Similarity 29.3%; Pred. No. 1.3e-08;
 Matches 81; Conservative 28; Mismatches 86; Indels 81; Gaps 14;

QY 19 GNPGLVNASKVEHDDROQ-LAAGSYSETFV-DLPSGTTAAHTPTPTTIPFAGH 76
 DB 18 GNPGLVNTDAGDLHDKDMQALAAEMDSETFVTGEPFDGAWPVR--IFTPAALIPFAGH 75
 QY 77 PTVGASWMLRER---GTPIN-LIOPAGIVGVSY-----HGDL 109
 DB 76 PTLGTAQVIRDLADGNPEVTVLDPGEVPEVBERDRETLMTQAPFEGQLAHD 135
 QY 110 LTAI-----SARSEWAPFAIHDDSL-----DALAAD-PADFPDDIAHYLMTWTD 155
 DB 136 LAAVAGLPADRLDDHPPEIVSTGLATIVPVPADNDAIGALDLDADAVV-----CD 189
 QY 156 RSA-----GSLRAMPFAAMGVTEBAGAAAIRITDYSR-----D 192
 DB 190 RDAKQVLAACPRSDNDLAVRFAPFYNVLEDPATGSSNCLAAVLAHRHMLSPAVE 249
 QY 193 LTTTQ---GKGSLIHTTWSPEG---WVRVGRVVS 221
 DB 250 ARVEQYEMGRPSLHLSTDSGEDISVRVGVGSVVA 285

RESULT 13
 Q4LTD6_9BURK PRELIMINARY; PRT; 286 AA.
 AC Q4LTD6_9
 DT 13-SEP-2005 (TREMBlrel. 31, Created)
 DT 13-SEP-2005 (TREMBlrel. 31, Last sequence update)
 DT 13-SEP-2005 (TREMBlrel. 31, Last annotation update)
 DE Phenazine biosynthesis PhzC/PhzF protein.
 GN ORFNames=Bcen2424DRAFT_4678;
 OS Burkholderia cenocepacia H12424.
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
 OC Burkholderiaceae; Burkholderia; Burkholderia cenocepacia complex.
 NCBI_TaxID=331272;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=H12424;
 RG US DOE Joint Genome Institute (JGI-PGF);
 RA Copeland A., Lucas S., Lapidus A., Barry K., Dettler C., Glavina T.,
 RA Hammon N., Istrati S., Plickus S., Richardson P.;
 RT "Sequencing of the draft genome assembly of Burkholderia cenocepacia
 H12424.";

RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=H12424;
 RA US DOE Joint Genome Institute (JGI-ORNL);
 RA Laitner F., Land M.;
 RT "Annotation of the draft genome assembly of Burkholderia cenocepacia
 H12424."
 RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
 CC -1- CAUTION: The sequence shown here is derived from an
 EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 preliminary data.
 DR EMBL; AAL01000017; EAM19398.1; -; Genomic DNA.
 CC SEQUENCE 286 AA; 30323 MW; 675BC2A8BBE7FAB5 CRC64;
 SQ
 Query Match 16.4%; Score 196; DB 2; Length 286;
 Best Local Similarity 31.0%; Pred. No. 2e-08;
 Matches 87; Conservative 32; Mismatches 90; Indels 72; Gaps 16;
 QY 7 VLKVFETDSDGNFNGPLGVI--NASKVEHRDRQQLAAGSYSETTFVDLPSPGTTAAH--T 63
 DB 8 LNVVFAESTFG-GNPLCVFPDARGMDATMQLAAGVFNSEITTFV-LP---SDRAHVR 62
 QY 64 IHTPRTEIPRAGPTVAGASWMLNERGTPIIN---TLQVPAGIVQSYHGDLTAISARSE 118
 DB 63 IFTPGYEMAFAGHPTTGTAHVRR---LNAAGALSLFRAGVVDVYARDVWTFIAPHA 119
 QY 119 WAEBEPAIHDL-----DSLDA-----LAAADADF-----PD----- 144
 DB 120 GMPNTAPCTLPDAOVAALLGLTEBDLLAPLMVNTGADOLLVAKDPAVRRAPDARV 179
 QY 145 DI-----AHYLTWTRDSAGSLRAMPAAVL--GTEDEATGAAIRTDYL---S 190
 DB 180 DIIPPSLSGRRTAVRPAFDPARPEIVSRFFPVKQSGVSEDPRTGSAACNLGGMILAA5 239
 QY 191 RD---LTTTQSGK---SLIHTTWSPEGWRVAVGRVSDG 223
 DB 240 HDLPALRVGEAGVGRPCLLHLSDVAAGKIGVGRTYELG 280

RESULT 14
 Q412T5_GIBZE PRELIMINARY; PRT; 309 AA.
 ID Q412T5;
 AC Q412T5;
 DT 13-SRP-2005 (TrEMBLrel. 31, Created)
 DT 13-SRP-2005 (TrEMBLrel. 31, last sequence update)
 DT 13-SRP-2005 (TrEMBLrel. 31, last annotation update)
 DE Hypothetical protein.
 GN ORFNames=FG08473.1;
 OS Glibberella zeae PH-1.
 OC Burkayota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 OC Hypocreomycetidae; Hypocreales; Nectriaceae; Glibberella.
 OX NCBI_TaxID=29533;
 OX 11)
 RN NUCLEOTIDE SEQUENCE.
 RC STRAIN=PH-1;
 RA Birren B., Nussbaum C., Abouelleil A., Allen N., Anderson S.,
 RA Atarchi H.M., Barma N., Baetien V., Bloom T., Boguslavsky L.,
 RA Boukagaler B., Butler J., Calvo S.B., Camarata J., Chang J.,
 RA Chougel Y., Collimore A., Cook A., Cooke P., Court B., Deatellano K.,
 RA Diaz J.S., Dodge S., Dooley K., Dorris L., Elkins T., Engels R.,
 RA Erickson J., Fero S., Ferreira P., Fitzgerald M., Gage D., Galagan J.,
 RA Gardyna S., Gnerre S., Graham L., Grand-Pierre N., Hafez N.,
 RA Hagopian D., Hagos B., Hall J., Horton L., Hulme W., Iliev I.,
 RA Jaffe D., Johnson R., Jones C., Kamel M., Kamet A., Karatas A.,
 RA Kaile C., Landers T., Levine R., Lindblad-Toh K., Liu G., Lui A.,
 RA Ma L.-J., Mabbitt R., Maclean C., Macdonald P., Major J., Manning J.,
 RA Matthews C., Mauceli E., McCarthy M., Meldrum J., Menes L.,
 RA Mihova T., Mlenga V., Murphy T., Naylor J., Nguyen C., Nicol R.,
 RA Nielsen C.B., Nordu C., O'Connor T., O'Donnell P., O'Neill D.,
 RA Oliver J., Peterson K., Punnhang P., Pierre N., Purcell S.,
 RA Rachupka A., Ramsamy U., Raymond C., Retta R., Rise C., Rogov P.,
 RA Roman J., Schauer S., Schuback R., Seaman S., Severy P., Smirnov S.,

RA Smith C., Spencer B., Stange-Thomann N., Stojanovic N., Stubbs M.,
 RA Talamas J., Teefaye S., Theodore J., Topham K., Travers M.,
 RA Vassiliev H., Venkataraman V.S., Viel R., Vo A., Wang S., Wilson B.,
 RA Wu X., Wyman D., Young G., Zainoun J., Zembek L., Zimmer A., Zody M.,
 RA Lander E.;
 RT "Fusarium graminearum genome sequence."
 RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
 CC -1- CAUTION: The sequence shown here is derived from an
 EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 preliminary data.
 DR EMBL; AACM01000342; EAA71290.1; -; Genomic DNA.
 CC Hypothetical protein.
 KW SEQUENCE 309 AA; 33637 MW; BB7CADB2CF373603 CRC64;
 SQ
 Query Match 16.4%; Score 196; DB 2; Length 309;
 Best Local Similarity 29.9%; Pred. No. 2.2e-08;
 Matches 78; Conservative 29; Mismatches 82; Indels 72; Gaps 13;
 QY 1 MAIEVSIVRYETDSDGNFNGPLGVI-----NASKVEHRDRQQLAAGSYSETTFVDLP5 54
 DB 1 MELPFTVLDFVETRTYR--GNPLAVVTIPADSNPKPTQEQKOTIAREFNISETVFIEHS 59
 QY 55 P--GSTTAAT--IHTPRTEIPRAGPTVAGASWMLNERGT-----PINTLOVPAGIV 102
 DB 60 PDADSDVTHRVVIDIFTSDEIIPRAGHPTTGAAVTLIPKGVDTVITAGPITALTQTRRGYI 119
 QY 103 Q-----VSYH-----GDLTAISARSEWAP-----FAIHDLDSL 132
 DB 120 QAAIPRNVKHSKTLADLS--PAPGQISPPAIRDCELRAPLSYNGMTFAIEPLD 178
 QY 133 ALAADPA-----DPPDD-----IAHYLTWTR-----SAGSLRAMPAAVLGVT 173
 DB 179 HIGRAISGVSLPVNLLDDDMRNGLLCKXYVNVDRREBDGYTVYSIRMAE---GAL 235
 QY 174 EDEATGAAIRTDYLSRDLT 194
 DB 236 EDPATGSAASALSHLSLOQT 256

RESULT 15
 Q8P1T5_XANAC PRELIMINARY; PRT; 292 AA.
 ID Q8P1T5;
 AC Q8P1T5;
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, last sequence update)
 DT 01-OCT-2002 (TrEMBLrel. 22, last annotation update)
 DE Hypothetical protein XAC2912.
 GN OrderedLocustNames=XAC2912;
 OS Xanthomonas axonopodis (pv. citri).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
 OC Xanthomonadaceae; Xanthomonas.
 OX NCBI_TaxID=92829;
 OX 11)
 RN NUCLEOTIDE SEQUENCE.
 RC STRAIN=306 / ATCC 13902 / XV 101;
 RX MEDLINE=22022145; PubMed=12024217; DOI=10.1038/417459a;
 RA da Silva A.C.R., Perro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
 RA Quagato R.B., Monteiro-Vitorello C.B., Van Sluys M.A.,
 RA Almeida N.F., Ut, Alves L.M.C., do Amaral A.M., Bertolini M.C.,
 RA Camargo L.B.A., Camarotte G., Canhaman P., Cardoso J., Chabergo F.,
 RA Ciapina L.P., Ciccarelli R.M.B., Coutinho L.L., Cursino-Santos J.R.,
 RA El-Dorri H., Faria J.B., Ferreira A.J.S., Ferreira R.C.C.,
 RA Perro M.I.T., Fomighieri B.F., Franco M.C., Greggio C.C., Gruber A.,
 RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
 RA Locali B.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
 RA Martins E.C., Melandus J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
 RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
 RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
 RA Spindola L.A.F., Takita M., Tamura R.B., Teixeira E.C., Tezza R.I.D.,
 RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
 RA Setubal J.C., Kitajima U.P.;
 RT "Comparison of the genomes of two Xanthomonas pathogens with differing
 host specificities.";

RL Nature 417:459-463(2002).
 DR EMBL; AE011933; AAM37157.1; -; Genomic_DNA.
 DR HSSP; P37757; 1SDJ.
 DR GO; GO:0003824; F:catalytic activity; IEA.
 DR GO; GO:0003058; P:biocynthesis; IEA.
 DR InterPro; IPR003719; PhzC_PhzF.
 DR Pfam; PF02567; PhzC_PhzF; 1.
 DR TIGRFAMs; TIGR00654; PhzF family; 1.
 DR Complete proteome; Hypothetical protein.
 SQ SEQUENCE 292 AA; 31038 MW; 2A9AADFE4816D976 CRC64;

Query Match 16.3%; Score 194; DB 2; Length 292;
 Best Local Similarity 29.1%; Pred. No. 3.1e-08;

Matches 83; Conservative 21; Mismatches 107; Indels 74; Gaps 11;

```
QY 8 LRVFTSDGNFGNPLGVI-NASKVEHRDRQQLAAGSGYSETFVLDLPSFGSTAAHATHT 66
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 9 LDVFAAGRAGS-GNPLGVFNAETLSSAQMOQIAAMTNTSETVFPPLPVSVDPADYHTRIFT 67
QY 67 PRTEIPFAGHPVTVGASMWLBERGT---PINTL--QVPAGIVQVSYHGDLTAISARSEWA 120
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 68 PRTEIPFAGHPVSYGAAVMAVTHGVTAYPEGRLRQCCAGVLPVDVFDRYGALLVRLRAP 127
QY 121 PEFALHDLDS-LDAL-----AAADPADF-----PDDIAHY 149
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 128 RARAIDTGYADLRPAACGLGASAOAPALMNNGNMMLBLADAAVQAAPDLAATA 187
QY 150 LWTWTRDSAG-----SLRARMFAANIQTEDERTGAATRTDYL----- 189
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 188 RUTQASAAACGLAIYAPMHDDADLVVRAFCPGDGIPEDPVTGSANACIAARLHGEDRLPG 247
QY 190 -----SRDLITQKGSLLHTTWSPEG--WVRVAGRVS DG 223
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 248 AGRVYASQGREY---GRDGRVHVVDNAGRWIGGATIQVTDG 288
```

Search completed: April 14, 2006, 17:32:13
 Job time : 88.1245 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: April 14, 2006, 17:32:59 ; Search time 21.0825 Seconds
(without alignments)
894.107 Million cell updates/sec

Title: US-10-620-246-66

Perfect score: 1192

Sequence: 1 MAIEVSVLRFVTDSDNFGN.....PEGWVAVGRVSDGVAQLD 228

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA:*
1: /cgn2_6/ptodata/1/1aa/5/ COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/6/ COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/H/ COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/PCTUS/ COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/RR/ COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/backtitle1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1192	100.0	228	2	US-09-050-739-66 Sequence 66, Appl
2	153.5	12.9	232	2	US-09-328-352-4558 Sequence 458, Ap
3	145.5	12.2	278	1	US-08-258-2618-20 Sequence 20, Appl
4	145.5	12.2	278	1	US-08-456-837-20 Sequence 20, Appl
5	145.5	12.2	278	1	US-08-457-342-20 Sequence 20, Appl
6	145.5	12.2	278	1	US-08-457-646A-20 Sequence 20, Appl
7	145.5	12.2	278	1	US-08-458-076A-20 Sequence 20, Appl
8	145.5	12.2	278	1	US-08-457-335A-20 Sequence 20, Appl
9	145.5	12.2	278	1	US-08-729-214-20 Sequence 20, Appl
10	145.5	12.2	278	2	US-09-028-934-20 Sequence 20, Appl
11	143.5	12.0	211	2	US-09-602-777A-242 Sequence 242, App
12	129	10.8	331	2	US-09-252-991A-21866 Sequence 21866, A
13	125	10.5	270	2	US-09-252-991A-20934 Sequence 20934, A
14	118	9.9	335	2	US-09-252-991A-27013 Sequence 27013, A
15	117.5	9.9	207	2	US-09-248-796A-16787 Sequence 16787, A
16	115	9.6	295	2	US-09-543-681A-6991 Sequence 6991, Ap
17	112.5	9.4	232	2	US-09-248-796A-16788 Sequence 16788, A
18	100.5	8.4	313	2	US-09-252-991A-20547 Sequence 20547, A
19	96.5	7.1	518	2	US-09-543-681A-6636 Sequence 6636, Ap
20	94.5	7.1	271	2	US-09-543-681A-4346 Sequence 4346, Ap
21	93	7.8	326	2	US-09-252-991A-22024 Sequence 22024, A
22	92	7.7	1482	2	US-09-410-551B-21 Sequence 21, Appl
23	92	7.7	1482	2	US-09-940-316B-21 Sequence 21, Appl
24	92	7.7	1488	2	US-09-410-551B-17 Sequence 17, Appl
25	92	7.7	1488	2	US-09-940-316B-17 Sequence 17, Appl
26	92	7.7	1509	2	US-09-410-551B-23 Sequence 23, Appl
27	92	7.7	1509	2	US-09-940-316B-23 Sequence 23, Appl

28	92	7.7	1517	2	US-09-410-551B-19 Sequence 19, Appl
29	92	7.7	1517	2	US-09-940-316B-19 Sequence 19, Appl
30	92	7.7	6396	2	US-09-410-551B-72 Sequence 72, Appl
31	92	7.7	6396	2	US-09-940-316B-72 Sequence 72, Appl
32	91	7.6	15281	1	US-08-471-119A-2 Sequence 2, Appl
33	90.5	7.6	286	2	US-09-134-000C-4899 Sequence 4899, Ap
34	89	7.5	3413	2	US-10-042-665A-8 Sequence 8, Appl
35	88	7.4	279	2	US-09-328-352-5581 Sequence 5581, Ap
36	87	7.3	197	2	US-09-605-703B-1974 Sequence 1974, Ap
37	86.5	7.3	1232	2	US-09-902-540-14264 Sequence 14264, A
38	86	7.2	285	2	US-09-328-352-4464 Sequence 4464, Ap
39	86	7.2	310	2	US-09-724-623-122 Sequence 122, App
40	86	7.2	613	2	US-09-248-796A-17277 Sequence 17277, A
41	86	7.2	1864	1	US-08-804-227C-3 Sequence 3, Appl
42	85.5	7.2	545	2	US-09-252-991A-21095 Sequence 21095, A
43	85	7.1	1253	1	US-07-920-281C-3 Sequence 3, Appl
44	85	7.1	1253	1	US-08-466-277-3 Sequence 3, Appl
45	85	7.1	1253	2	US-09-688-842-3 Sequence 3, Appl

ALIGNMENTS

RESULT 1
US-09-050-739-66
Sequence 66, Application US/09050739
Patent No. 6641814
GENERAL INFORMATION:
APPLICANT: ANDERSEN, Peter
APPLICANT: NIELSEN, Rikke
APPLICANT: OETTINGER, Thomas
APPLICANT: RASMUSSEN, Peter Blak
APPLICANT: ROSENKRANDS, Ida
APPLICANT: WELDINGH, Karin
APPLICANT: FLORIO, Walter
TITLE OF INVENTION: NUCLEIC ACIDS FRAGMENTS AND POLYPEPTIDE FRAGMENTS
FILE REFERENCE: 670001-2002.1
CURRENT APPLICATION NUMBER: US/09/050,739
CURRENT FILING DATE: 1998-03-30
EARLIER APPLICATION NUMBER: 0376/97
EARLIER FILING DATE: 1997-04-02
EARLIER APPLICATION NUMBER: 1277/97
EARLIER FILING DATE: 1997-11-10
EARLIER APPLICATION NUMBER: 60/044,624
EARLIER FILING DATE: 1997-04-18
EARLIER APPLICATION NUMBER: 60/070,488
NUMBER OF SEQ ID NOS: 173
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 66
LENGTH: 228
TYPE: PRT
ORGANISM: Mycobacterium tuberculosis
US-09-050-739-66

Query Match 100.0%; Score 1192; DB 2; Length 228;
Best Local Similarity 100.0%; Pred. No. 1, 1e-121;
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MAIEVSVLRFVTDSDNFGNPLGVINASKVHRDROQLAAQSGYSEITFVLDPSPGSTTA	60
DB	1	MAIEVSVLRFVTDSDNFGNPLGVINASKVHRDROQLAAQSGYSEITFVLDPSPGSTTA	60
QY	61	HATHTPRKTRIPAGHPYTGASWMLBKRGPINTLOVPAGIVVSTHGDTTASASSEWA	120
DB	61	HATHTPRKTRIPAGHPYTGASWMLBKRGPINTLOVPAGIVVSTHGDTTASASSEWA	120
QY	121	PEFAHIDLSLDLAAADPPDDIAHYLMTWTDSDSAGSLRAMPANGVTEDEATGA	180
DB	121	PEFAHIDLSLDLAAADPPDDIAHYLMTWTDSDSAGSLRAMPANGVTEDEATGA	180
QY	181	AAIRTDYLSRDYLTITQKGLIHTTWSPEGWVAVGRVSDGVAQLD	228

FILING DATE: 01-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/457,205
FILING DATE: 01-JUN-1995
APPLICATION NUMBER: 08/258,261
FILING DATE: 08-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Elmer, James Scott
REGISTRATION NUMBER: 36,129
REFERENCE/DOCKET NUMBER: CGC 1506/CIP3
TELEPHONE: 919-541-8614
TELEFAX: 919-541-8689
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 278 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULAR TYPE: protein
US-08-456-837-20

Query Match 12.2%; Score 145.5; DB 1; Length 278;

Best Local Similarity 28.3%; Pred. No. 2e-07; Mismatches 83; Indels 41; Gaps 9;

Matches 62; Conservative 33; Mismatches 83; Indels 41; Gaps 9;

QY 19 GNPLGV-INASKVHRDRQOLAOSGYSETIPVDLPSPGSTTAATHTPTREIPFAGHP 77
17 GNVAVFPDADDSABQMORIAEMNLSETTFV-LKRNCGDALIRLFTVNEBLPFAGHP 75
QY 78 TVGASWMLBERGTPINTLQVPAGIVQVSY-----HGDLTAISARSEWAPFAI-HDDSL 131
Db 76 LGSTDLALGAR-TDNHRLPLETQMGTAFLERQNGSVIAASMDQPIPTWTALGRDAEL 134
QY 132 DALAADPADPPDDIAH-----YMTWTRDS---AGS--- 160
Db 135 KALGIDB-STFPPIEYHNGPRHVFGVPSIALSALHPDHRALYSFHDMAINCFAAGRR 193
QY 161 LBARMPAANLGVTEDEATGAATIRITDYLSDRLTTIOGK 199
Db 194 WSRMPSPAYGVVEDAXTGSNAGPLAIHLARKQIERFGQ 232

RESULT 5

US-08-457-342-20

Sequence 20, Application US/08457342

Patent No. 5662898

GENERAL INFORMATION:

APPLICANT: Schupp, Thomas

APPLICANT: Ligon, James M.

APPLICANT: Beck, James Joseph

APPLICANT: Hill, Dwight Steven

APPLICANT: Ryals, John Andrew

APPLICANT: Gaffney, Thomas Deane

APPLICANT: Lam, Stephen Ting

APPLICANT: Hammer, Phillip B.

APPLICANT: Uknes, Scott Joseph

TITLE OF INVENTION: Genes for the syntheses of

TITLE OF INVENTION: antipathogenic substances

NUMBER OF SEQUENCES: 22

CORRESPONDENCE ADDRESS:

ADDRESSEE: Ciba-Geigy Corporation

STREET: 7 Skyline Drive

CITY: Hawthorne

STATE: NY

COUNTRY: USA

ZIP: 10532

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

APPLICATION NUMBER: US/08/457,342
FILING DATE: 01-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/457,205
FILING DATE: 01-JUN-1995
APPLICATION NUMBER: 08/258,261
FILING DATE: 08-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Elmer, James Scott
REGISTRATION NUMBER: 36,129
REFERENCE/DOCKET NUMBER: CGC 1506/CIP3
TELEPHONE: 919-541-8614
TELEFAX: 919-541-8689
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 278 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULAR TYPE: protein
US-08-457-342-20

Query Match 12.2%; Score 145.5; DB 1; Length 278;

Best Local Similarity 28.3%; Pred. No. 2e-07; Mismatches 83; Indels 41; Gaps 9;

Matches 62; Conservative 33; Mismatches 83; Indels 41; Gaps 9;

QY 19 GNPLGV-INASKVHRDRQOLAOSGYSETIPVDLPSPGSTTAATHTPTREIPFAGHP 77
17 GNVAVFPDADDSABQMORIAEMNLSETTFV-LKRNCGDALIRLFTVNEBLPFAGHP 75
QY 78 TVGASWMLBERGTPINTLQVPAGIVQVSY-----HGDLTAISARSEWAPFAI-HDDSL 131
Db 76 LGSTDLALGAR-TDNHRLPLETQMGTAFLERQNGSVIAASMDQPIPTWTALGRDAEL 134
QY 132 DALAADPADPPDDIAH-----YMTWTRDS---AGS--- 160
Db 135 KALGIDB-STFPPIEYHNGPRHVFGVPSIALSALHPDHRALYSFHDMAINCFAAGRR 193
QY 161 LBARMPAANLGVTEDEATGAATIRITDYLSDRLTTIOGK 199
Db 194 WSRMPSPAYGVVEDAXTGSNAGPLAIHLARKQIERFGQ 232

RESULT 6

US-08-457-646A-20

Sequence 20, Application US/08457646A

Patent No. 5679560

GENERAL INFORMATION:

APPLICANT: Schupp, Thomas

APPLICANT: Ligon, James M.

APPLICANT: Beck, James Joseph

APPLICANT: Hill, Dwight Steven

APPLICANT: Ryals, John Andrew

APPLICANT: Gaffney, Thomas Deane

APPLICANT: Lam, Stephen Ting

APPLICANT: Hammer, Phillip B.

APPLICANT: Uknes, Scott Joseph

TITLE OF INVENTION: Genes for the syntheses of

TITLE OF INVENTION: antipathogenic substances

NUMBER OF SEQUENCES: 22

CORRESPONDENCE ADDRESS:

ADDRESSEE: Ciba-Geigy Corporation

STREET: 7 Skyline Drive

CITY: Hawthorne

STATE: NY

COUNTRY: USA

ZIP: 10532

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

```
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/457,646A
/ FILING DATE: 01-JUN-1995
/ CLASSIFICATION: 530
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/457,205
/ FILING DATE: 01-JUN-1995
/ APPLICATION NUMBER: 08/258,261
/ FILING DATE: 08-JUN-1994
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Bimler, James Scott
/ REGISTRATION NUMBER: 36,129
/ REFERENCE/DOCKET NUMBER: CGC 1506/CIP3
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 919-541-8614
/ TELEFAX: 919-541-8689
/ INFORMATION FOR SEQ ID NO: 20:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 278 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
US-08-457-646A-20
```

```
Query Match 12.2%; Score 145.5; DB 1; Length 278;
Best Local Similarity 28.3%; Pred. No. 2e-07;
Matches 62; Conservative 33; Mismatches 83; Indels 41; Gaps 9;
```

```
QY 19 GNPPLGV-YNASKVEHRDQGLAAGSGSETIFVDLPSPGSTTAHATHTPRTIPPAGHP 77
DB 17 GNPVAVFPDDDISAQMQRIRAKENMLSETTV-LKPRNGDALLIRIFPVNELPPAGHP 75
QY 78 TVGASWMLERGTPIINTLOVPAGIVQVSY-----HGDLTAISARSEWAEPAI-HDLDSL 131
DB 76 LGDTIALGAR-TDNRRLFLFTQMGTLAPELERQNGSVTAASMDQIPPTWALGRDAELL 134
QY 132 DALAADPADPDDIAH-----YLTWTDS---AGS--- 160
DB 135 KALGISD-STFPIEITYHNGPRHVFVGLPSIALSALHPDHRALYSFHDVAINCFAAGRR 193
QY 161 LRBPMPANLGVTEDEATGAARITRDYLSRDLTTQOK 199
DB 194 WRSRMFSPAYGVVEDAKTGSAAGPLAIHLARHGQIFGQ 232
```

RESULT 7

```
US-08-458-076A-20
/ Sequence 20, Application US/08458076A
/ Patent No. 5698425
/ GENERAL INFORMATION:
/ APPLICANT: Schupp, Thomas
/ APPLICANT: Ligon, James M.
/ APPLICANT: Beck, James Joseph
/ APPLICANT: Hill, Dwight Steven
/ APPLICANT: Ryals, John Andrew
/ APPLICANT: Gaffney, Thomas Deane
/ APPLICANT: Lam, Stephen Ting
/ APPLICANT: Hammer, Phillip E.
/ APPLICANT: Uknes, Scott Joseph
/ TITLE OF INVENTION: Genes for the synthesis of
/ TITLE OF INVENTION: antipathogenic substances
/ NUMBER OF SEQUENCES: 22
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Ciba-Geigy Corporation
/ STREET: 7 Skyline Drive
/ CITY: Hawthorne
/ STATE: NY
/ COUNTRY: USA
/ ZIP: 10532
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: floppy disk
/ OPERATING SYSTEM: PC-DOS/MS-DOS
```

```
/ SOFTWARE: Patent Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/458,076A
/ FILING DATE: 01-JUN-1995
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/457,205
/ FILING DATE: 01-JUN-1995
/ APPLICATION NUMBER: 08/258,261
/ FILING DATE: 08-JUN-1994
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Bimler, James Scott
/ REGISTRATION NUMBER: 36,129
/ REFERENCE/DOCKET NUMBER: CGC 1506/CIP3
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 919-541-8614
/ TELEFAX: 919-541-8689
/ INFORMATION FOR SEQ ID NO: 20:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 278 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
US-08-458-076A-20
```

```
Query Match 12.2%; Score 145.5; DB 1; Length 278;
Best Local Similarity 28.3%; Pred. No. 2e-07;
Matches 62; Conservative 33; Mismatches 83; Indels 41; Gaps 9;
```

```
QY 19 GNPPLGV-YNASKVEHRDQGLAAGSGSETIFVDLPSPGSTTAHATHTPRTIPPAGHP 77
DB 17 GNPVAVFPDDDISAQMQRIRAKENMLSETTV-LKPRNGDALLIRIFPVNELPPAGHP 75
QY 78 TVGASWMLERGTPIINTLOVPAGIVQVSY-----HGDLTAISARSEWAEPAI-HDLDSL 131
DB 76 LGDTIALGAR-TDNRRLFLFTQMGTLAPELERQNGSVTAASMDQIPPTWALGRDAELL 134
QY 132 DALAADPADPDDIAH-----YLTWTDS---AGS--- 160
DB 135 KALGISD-STFPIEITYHNGPRHVFVGLPSIALSALHPDHRALYSFHDVAINCFAAGRR 193
QY 161 LRBPMPANLGVTEDEATGAARITRDYLSRDLTTQOK 199
DB 194 WRSRMFSPAYGVVEDAKTGSAAGPLAIHLARHGQIFGQ 232
```

RESULT 8

```
US-08-457-335A-20
/ Sequence 20, Application US/08457335A
/ Patent No. 5723759
/ GENERAL INFORMATION:
/ APPLICANT: Schupp, Thomas
/ APPLICANT: Ligon, James M.
/ APPLICANT: Beck, James Joseph
/ APPLICANT: Hill, Dwight Steven
/ APPLICANT: Ryals, John Andrew
/ APPLICANT: Gaffney, Thomas Deane
/ APPLICANT: Lam, Stephen Ting
/ APPLICANT: Hammer, Phillip E.
/ APPLICANT: Uknes, Scott Joseph
/ TITLE OF INVENTION: Genes for the synthesis of
/ TITLE OF INVENTION: antipathogenic substances
/ NUMBER OF SEQUENCES: 22
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Ciba-Geigy Corporation
/ STREET: 7 Skyline Drive
/ CITY: Hawthorne
/ STATE: NY
/ COUNTRY: USA
/ ZIP: 10532
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: floppy disk
/ OPERATING SYSTEM: IBM PC compatible
```


; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 21986
 ; LENGTH: 331
 ; TYPE: PR1
 ; ORGANISM: Pseudomonas aeruginosa
 US-09-252-991A-21986

Query Match 10.8%; Score 129; DB 2; Length 331;
 Best Local Similarity 27.6%; Pred. No. 1.6e-05;
 Matches 51; Conservative 24; Mismatches 76; Indels 32; Gaps 10;

QY 19 GNPGLVYNASKVEHRDQOLAAGSYSETIFVDLPSPGTTAHTHTPTETIPFAGHP 77
 DB 85 GNAVAAYIGADRLSSRQMLFPAATQLETTFLRPTVAENDYVRIFTPLEELPFAGHP 144
 QY 78 TVGA-SWTLBERG---TPINTLOVPAGIVQSYHGLDITAISSBEMPER---AIND-- 127
 DB 145 TLGSCQVWNLQGGGNADEIVQBCLAGLIRRRKALLSFA-----APLLRGGADEV 199
 QY 128 LDSLALAAADPADFPDDIAHYLTWTDRSAGSLRAMPANLGVTEDEATGAARITD 187
 DB 200 LRRIESELGSPQVRS-----QWVDNGRGVAVAL-----ATIDE---VLAIR-PD 243
 QY 188 YLSRD 192
 DB 244 YAKLD 248

RESULT 13
 US-09-252-991A-20934
 ; Sequence 20934, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252,991A
 ; PRIOR FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 20934
 ; LENGTH: 270
 ; TYPE: PR1
 ; ORGANISM: Pseudomonas aeruginosa
 US-09-252-991A-20934

Query Match 10.5%; Score 125; DB 2; Length 270;
 Best Local Similarity 26.4%; Pred. No. 3.2e-05;
 Matches 70; Conservative 33; Mismatches 108; Indels 54; Gaps 13;

QY 3 IEVSVLRVFTDSGNF-GNPGLVYNASK-VEHRDQOLAAGSYSETIFVDLPSPGTTA 60
 DB 12 MELTITQVDAFASPGNPAVCPDLAMTDERLQIAENNLSEAFV-VGRDD--Y 68
 QY 61 HATHTPRETIPFAGHPVTGASWMLRER---GTPINTLOVPAGIVQSYHGLDITAISSARS 117
 DB 69 RLKMFPTQVEVDLCGHATTATATWLTIHKLDASPVLFATRSGBLSVRSGDSLAMDFFA 128
 QY 118 EMAPERAIHDLSDLAADPADFPDDIAHYLTWTDR----- 156
 DB 129 K-RPEPCATPDGLLEAGIAEAVLKTDDYLVVVDDEKTIALAADPFARLKLPGKV 185
 QY 157 --SAGSLR---ARMFANLGVTEDEATGAARITDYLGRDLTIT-----QG---KSL 202
 DB 186 AVTARSGRPFVSRWGPVNGVNEDEPTYSHTSIAPYMAQRGLKRLSLAEGGARKRL 245
 QY 203 IHTTWSPEGWVRVAGRVSDGVAQL 227

DB 246 -----ECVVR-GERVVISGKAL 262

RESULT 14
 US-09-252-991A-27013
 ; Sequence 27013, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252,991A
 ; PRIOR FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 27013
 ; LENGTH: 335
 ; TYPE: PR1
 ; ORGANISM: Pseudomonas aeruginosa
 US-09-252-991A-27013

Query Match 9.9%; Score 118; DB 2; Length 335;
 Best Local Similarity 23.2%; Pred. No. 0.00026;
 Matches 61; Conservative 34; Mismatches 98; Indels 70; Gaps 10;

QY 19 GNPGLVYNASKVEHRDQOLAAGSYSETIFVDLPSPGTTAHTHTPTETIPFAGHP 77
 DB 69 GNGIAVFDASTLDDAQMAMTRELRFESIFL-LGDDPRAPRAIRFTLEELPFAGHP 127
 QY 78 TVGAS-----WTLBERGTPINTLOVPAGIVQSYHGLD-----TAIS 114
 DB 128 LLSAALLHLRLGQDNEQHHTLHLASKSVLRSG---SGFYAENDQGRAFGATPDA 184
 QY 115 ARSEM-APERAIHDLSDLAADPADFPDDIAHYLTWTDRSAGSLRA-----R 164
 DB 185 GTCRWABEASL-----SANDLSGHPRRVYSTGLPYLLPTAALGRARQVNDLQERLXD 240
 QY 165 MFAA-----NLGVTEDEATGAARITDYLSDLTITQKSLIH----- 204
 DB 241 LGAAPYLLVDVDRERGTWDSGLVBDVATGSAAGPVAAYLVGYLAARGPVPLHGRF 300
 QY 205 -----TWSPEGWVRVAGRV 219
 DB 301 LERSRLDVQVATDGSVRVGHV 323

RESULT 15
 US-09-248-796A-16787
 ; Sequence 16787, Application US/09248796A
 ; Patent No. 6747137

; GENERAL INFORMATION:
 ; APPLICANT: Keith Weinstock et al.
 ; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 107196.132
 ; CURRENT APPLICATION NUMBER: US/09/248,796A
 ; PRIOR FILING DATE: 1999-02-12
 ; PRIOR APPLICATION NUMBER: US 60/074,725
 ; PRIOR FILING DATE: 1998-02-13
 ; PRIOR APPLICATION NUMBER: US 60/096,409
 ; PRIOR FILING DATE: 1998-08-13
 ; NUMBER OF SEQ ID NOS: 28208
 ; SEQ ID NO 16787
 ; LENGTH: 207
 ; TYPE: PR1
 ; ORGANISM: Candida albicans
 US-09-248-796A-16787

Query Match	9.9%	Score 117.5;	DB 2;	Length 207;
Best Local Similarity	28.2%;	Pred. No. 0.00014;		
Matches 46;	Conservative 31;	Mismatches 53;	Indels 33;	Gaps 9;

```

QY 10 PTTSDGNGNPLGY--NASKVEHNRQOLAQSGSEITFYVLLPBGSTTAATH---65
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 49 VFTNK-YLGNVAIYDSDNITTTOMCKIARTNLSETTEI--LTPKSIAYSRIEFT 105
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY 66 TPRTETPAGHPYGVASWMLRERG--TPINTLVQV---PAGIVQSVGHDLTAISARSE-118
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 106 SGNELPRAGHTLTGTAPLLEDGKIKPNDNGOIIQECGAGLVKISV--EKTNNNSNEL 163
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY 119 -----WAFPAIHDIHSLDALAADPADPDDIAHYLTWTWD 155
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 164 PPLLSFQLPYFKENHD-----DKVIELHLSMNGTID 195

```

Search completed: April 14, 2006, 17:37:53
Job time : 22.0825 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using bw model

Run on: April 14, 2006, 18:27:13 ; Search time 80.0108 seconds
(without alignments)
1190.652 Million cell updates/sec

Title: US-10-620-246-66

Perfect score: 1192
Sequence: 1 MAIEVSVLRFTDSGNFGN.....PEGWVAVGRVSDVAQLD 228

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA Main: *
1: /cgn2_6/prodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/prodata/1/pubpaa/US08_PUBCOMB.pep:*
3: /cgn2_6/prodata/1/pubpaa/US09_PUBCOMB.pep:*
4: /cgn2_6/prodata/1/pubpaa/US10_PUBCOMB.pep:*
5: /cgn2_6/prodata/1/pubpaa/US10B_PUBCOMB.pep:*
6: /cgn2_6/prodata/1/pubpaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1192	100.0	228	3 US-09-791-171-66	Sequence 66, Appl
2	1192	100.0	228	3 US-09-804-980-66	Sequence 66, Appl
3	1192	100.0	228	4 US-10-620-246-66	Sequence 66, Appl
4	427	35.8	214	4 US-10-156-761-13456	Sequence 13456, A
5	179.5	15.1	306	4 US-10-425-115-210632	Sequence 210632, A
6	159.5	13.4	280	3 US-09-738-626-6518	Sequence 6518, Ap
7	151.5	12.7	278	4 US-10-389-647-621	Sequence 621, App
8	128	10.7	272	4 US-10-156-761-14521	Sequence 14521, A
9	125	10.5	120	4 US-10-425-115-230986	Sequence 230986, A
10	112.5	9.4	274	4 US-10-166-087-30	Sequence 30, Appl
11	110	9.2	288	5 US-10-733-869A-87	Sequence 87, Appl
12	110	9.2	375	5 US-10-450-763-52819	Sequence 52819, A
13	100.5	8.4	636	4 US-10-156-761-13432	Sequence 13432, A
14	100	8.4	214	4 US-10-424-599-178493	Sequence 178493, A
15	99	8.3	217	5 US-10-450-763-52821	Sequence 52821, A
16	99	8.3	284	4 US-10-369-493-23278	Sequence 23278, A
17	99	8.3	301	4 US-10-424-599-200127	Sequence 200127, A
18	98.5	8.3	775	4 US-10-156-761-10241	Sequence 10241, A
19	96.5	8.1	513	4 US-10-282-122A-69017	Sequence 69017, A
20	95	8.0	1107	4 US-10-282-122A-59131	Sequence 59131, A
21	94	7.9	279	4 US-10-369-493-11888	Sequence 11888, A
22	92.5	7.8	3808	4 US-10-378-083-7	Sequence 7, Appl
23	92.5	7.8	5674	5 US-10-844-716-10	Sequence 10, Appl
24	92.5	7.7	6043	5 US-10-369-493-23012	Sequence 23012, A
25	92	7.7	281	4 US-10-425-115-227405	Sequence 227405, A
26	92	7.7	296	4 US-10-425-115-227405	Sequence 227405, A
27	92	7.7	315	4 US-10-425-114-43651	Sequence 43651, A

28	92	7.7	315	4 US-10-425-114-47781	Sequence 47781, A
29	92	7.7	327	4 US-10-369-493-12173	Sequence 12173, A
30	92	7.7	1482	3 US-09-940-3168-21	Sequence 21, Appl
31	92	7.7	1488	3 US-09-940-3168-17	Sequence 17, Appl
32	92	7.7	1509	3 US-09-940-3168-23	Sequence 23, Appl
33	92	7.7	1517	3 US-09-940-3168-19	Sequence 19, Appl
34	92	7.7	6396	3 US-09-940-3168-72	Sequence 72, Appl
35	90.5	7.6	238	4 US-10-425-115-248384	Sequence 248384, A
36	90	7.6	1642	5 US-10-741-600-1176	Sequence 1176, Ap
37	90	7.6	1642	5 US-10-741-600-1176	Sequence 1176, Ap
38	90	7.6	3024	4 US-10-408-765A-1977	Sequence 1977, Ap
39	90	7.6	3396	4 US-10-788-792-170	Sequence 170, App
40	90	7.6	3396	5 US-10-741-600-1172	Sequence 1172, Ap
41	90	7.6	3396	5 US-10-741-600-1173	Sequence 1173, Ap
42	90	7.6	3396	5 US-10-631-467-773	Sequence 773, App
43	89	7.5	3413	5 US-10-042-665A-8	Sequence 8, Appl
44	88.5	7.4	315	4 US-10-425-114-66693	Sequence 66693, A
45	88.5	7.4	899	4 US-10-203-295-30	Sequence 30, Appl

ALIGNMENTS

```
RESULT 1
US-09-791-171-66
; Sequence 66, Application US/09791171
; Patent No. US20020094336A1
; GENERAL INFORMATION:
; APPLICANT: ANDERSEN, Peter
; APPLICANT: NIELSEN, Kikke
; APPLICANT: OSTINGER, Thomas
; APPLICANT: RASMUSSEN, Peter Birk
; APPLICANT: ROSENKRANDS, Ida
; APPLICANT: WELDMING, Karin
; APPLICANT: FIORIO, Walter
; TITLE OF INVENTION: NUCLEIC ACIDS FRAGMENTS AND POLYPEPTIDE FRAGMENTS
; FILE REFERENCE: 670001-2002.1
; CURRENT APPLICATION NUMBER: US/09/791,171
; CURRENT FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 09/050,739
; PRIOR FILING DATE: 1998-03-30
; PRIOR APPLICATION NUMBER: 0376/97
; PRIOR FILING DATE: 1997-04-02
; PRIOR APPLICATION NUMBER: 60/070,488
; PRIOR FILING DATE: 1998-01-05
; NUMBER OF SEQ ID NOS: 173
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 66
; LENGTH: 228
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-09-791-171-66
Query Match 100.0%; Score 1192; DB 3; Length 228;
Best local similarity 100.0%; Pred. No. 1.8e-116;
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Cy 1 MAIEVSVLRFTDSGNFGNPGVIVASVYEHDRQQLAQSISFTIFVLDPSGSTA 60
Db 1 MAIEVSVLRFTDSGNFGNPGVIVASVYEHDRQQLAQSISFTIFVLDPSGSTA 60
Cy 61 HATIHPTREIPAGHPVYASVWTLRERGPINTLQVPAGIVQVSYHGDLTASARSEWA 120
Db 61 HATIHPTREIPAGHPVYASVWTLRERGPINTLQVPAGIVQVSYHGDLTASARSEWA 120
Cy 121 PEFALHDLSDLAALADPADPDDIAHYLWTTDRSAGSLRARMFANIGVTEDEATCA 180
Db 121 PEFALHDLSDLAALADPADPDDIAHYLWTTDRSAGSLRARMFANIGVTEDEATCA 180
```

```
OY      181 AAIRITDYSRDLTITGKSLIHHTWSPSGMVRVAGRVSDGAQLD 228
        |||||
DB       181 AAIRITDYSRDLTITGKSLIHHTWSPSGMVRVAGRVSDGAQLD 228
```

```

US-09-804-980-66
RESULT 2
: Sequence 66, Application US/09804980
: Publication No. US20030147897A1
: GENERAL INFORMATION:
: APPLICANT: Statens Serum Institut
: APPLICANT: Andersen, Peter
: TITLE OF INVENTION: M. tuberculosis Antigen
: FILE REFERENCE: 670001-2002.4
: CURRENT APPLICATION NUMBER: US/09/804,980
: CURRENT FILING DATE: 2001-03-12
: NUMBER OF SEQ ID NOS: 257
: SOFTWARE: PatentIn version 3.0
SEQ ID NO 66
LENGTH: 228
TYPE: PRN
: ORGANISM: Mycobacterium tuberculosis
US-09-804-980-66

```

Query Match	100.0%;	Score 1192;	DB 3;	Length 228;
Best Local Similarity	100.0%;	Pred. No. 1.8e-116;		
Matches 228;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0

QY	1	MALEVSALRYFTSDSDGNFGNPLGVINASVTEHRDQOOLAAOSGETTFIDLPPSGTTA	60
QY	1	MALEVSALRYFTSDSDGNFGNPLGVINASVTEHRDQOOLAAOSGETTFIDLPPSGTTA	60
Db	1	MALEVSALRYFTSDSDGNFGNPLGVINASVTEHRDQOOLAAOSGETTFIDLPPSGTTA	60
QY	61	HATHTPRTEIPRAPHPTVGASWMLERGTPINTLQVAGIQVSYHGDLTALISARSEMA	120
Db	61	HATHTPRTEIPRAPHPTVGASWMLERGTPINTLQVAGIQVSYHGDLTALISARSEMA	120
QY	121	PERAIHDLDSLDAALAADPADPPDDIAHYLTMTWDRSAGSLRARMFAANLCVTEDEATGA	180
Db	121	PERAIHDLDSLDAALAADPADPPDDIAHYLTMTWDRSAGSLRARMFAANLCVTEDEATGA	180
QY	181	AAARIDVLSRDLTTIQKGSLLHTTWSPEGVVRVAGVVSDDVQAQD	228
Db	181	AAARIDVLSRDLTTIQKGSLLHTTWSPEGVVRVAGVVSDDVQAQD	228

```

RESULT 3
US-10-620-246-66
/ Sequence 66, Application US/10620246
/ Publication No. US20040115211A1
/ GENERAL INFORMATION:
/ APPLICANT: ANDERSEN, Peter
/ APPLICANT: NIELSEN, Rikke
/ APPLICANT: OETTINGER, Thomas
/ APPLICANT: RASMUSSEN, Peter Birk
/ APPLICANT: ROSENKRANDS, Ida
/ APPLICANT: WELDINGH, Karin
/ APPLICANT: FLORIO, Walter
/ TITLE OR INVENTION: NUCLEIC ACIDS FRAGMENTS AND POLYPEPTIDE FRAGMENTS
/ TITLE OR INVENTION: DERIVED FROM M. TUBERCULOSIS
/ FILE REFERENCE: 670001-2062.1A
/ CURRENT APPLICATION NUMBER: US/10/620,246
/ CURRENT FILING DATE: 2003-07-15
/ PRIOR APPLICATION NUMBER: 09/050,739
/ PRIOR FILING DATE: 1998-03-30
/ PRIOR APPLICATION NUMBER: 0376/97
/ PRIOR FILING DATE: 1997-04-02
/ PRIOR APPLICATION NUMBER: 1277/97
/ PRIOR FILING DATE: 1997-11-10
/ PRIOR APPLICATION NUMBER: 60/044,624
/ PRIOR FILING DATE: 1997-04-18
/ PRIOR APPLICATION NUMBER: 60/070,488
/ PRIOR FILING DATE: 1998-01-05

```

```

? PRIOR APPLICATION NUMBER: 10/138,473
? PRIOR FILING DATE: 2002-05-02
? PRIOR APPLICATION NUMBER: 09/791,171
? PRIOR FILING DATE: 2001-02-20
? PRIOR APPLICATION NUMBER: 09/415,884
? PRIOR FILING DATE: 1999-10-08
? PRIOR APPLICATION NUMBER: 60/116,673
? PRIOR FILING DATE: 1999-01-21
? PRIOR APPLICATION NUMBER: 1281/98
? PRIOR FILING DATE: 1998-10-08
? NUMBER OF SEQ ID NOS: 173
? SOFTWARE: PatentIn Ver. 2.0
? SEQ ID NO 66
? LENGTH: 228
? TYPE: PRT
? ORGANISM: Mycobacterium tuberculosis
US-10-620-246-66

Query Match          100.0%; Score 1192; DB 4; Length 228;
Best Local Similarity 100.0%; Pred. No. 1.8e-116;
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0

```

[illegible]

```

RESULT 4
US-10-156-761-13456
? Sequence 13456, Application US/10156761
? Publication No. US20030119018A1
? GENERAL INFORMATION:
? APPLICANT: OMURA, SATOSHI
? APPLICANT: IKEDA, HARUO
? APPLICANT: ISHIKAWA, JUN
? APPLICANT: HORIKAWA, HIROSHI
? APPLICANT: SHIBA, TADAYOSHI
? APPLICANT: SAKAKI, YOSHIYUKI
? APPLICANT: HATTORI, MASAHIRA
? TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
? FILE REFERENCE: 245-262
? CURRENT APPLICATION NUMBER: US/10/156, 761
? PRIORITY FILING DATE: 2002-05-29
? PRIOR APPLICATION NUMBER: JP 2001-204089
? PRIOR FILING DATE: 2001-05-30
? PRIOR APPLICATION NUMBER: JP 2001-272697
? NUMBER OF SEQ ID NOS: 08-02
? SEQ ID NO 13456
? LENGTH: 214
? TYPE: PRN
? ORGANISM: Streptomyces avermitilis
US-10-156-761-13456

```

Query March 35.8% Score 427; DB 4; length 214;
Best local similarity 43.9%; Pred. No. 3, 66-36;
Matches 98; Conservative 26; Mismatches 79; Indels 20; Gaps 6

4 EVSLVLRFTDSDGNGFNPLGYINASKV--EHRDPOOLAAGSYSETTIVDLPSFGSTAH 61
3 DYDLRLRFGCGGGGNGELGVRRGSLVPEACADQALAAKKIGSETTIVEVDLPPERYV--- 58
b


```

QY 62 ATHTPTETIPKPGHPTVGSWTLRENGTPIINTLQVAGIVQSYHDDLTAISRSWAP 121
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 59 IDYTPTLRLPFAGHPCVGTAMLL-----DVPELVTPAGVGTRLDEPFWIRARPEMAP 113
QY 122 EFAIHDLDS---IDALAAADPADPDDIAHYLMTWTWTSRSGSLRABMF--AANLGVTEDEA 177
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 114 PRTLRQYATAAEVDALFVPPGGEV-----YANAMEDBAAGRYARAFPGRDGIEDEEA 168
QY 178 TGAATAIRITDLSRDLTITQKSLIHTTWSPEGWVAVAGRNV 220
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 169 TGAALLLTDRLRGALNITIGRSGSLITAPQAGWVAVGRVYL 211

```

```

RESULT 5
US-10-425-115-210632
; Sequence 210632, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 210632
; LENGTH: 306
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(306)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_123694C.1.pep
US-10-425-115-210632

```

```

Query Match      15.1%; Score 179.5; DB 4; Length 306;
Best Local Similarity 30.5%; Pred. No. 5.8e-10;
Matches 83; Conservative 30; Mismatches 80; Indels 79; Gaps 16;

QY 1 MALEVSIVRYETSDGNF--GNPLGV--INAS--KVEHRDROQLAAGSGSETIFV--DLPs 54
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 MELPYTLADVFTRT--RFAGNPUGVRTIPASGPKPTQAKOSIRKRFNFSEITFIHIDVAD 58
QY 55 PG-STTAHATITHTPTEIPKPGHPTVGSWTLRENG-----TPINTLQVPAG 100
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 59 PATNTTRITDIPITLTAIEIPKPGHPTIGAVTLISQGVTOITKAGLIPVPTGPDVSIG 118
QY 101 IVO-----VSYHG-DLTAISARSEW-----APRPAL-----HIDSLDLA 135
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 119 IPHNVHLHATSLASYPGLBPAHLQRNBITROLELAAPVSLVGMTEFVEIPLSEILLA 178
QY 136 AADPA--DEPDD-----IANYLMTWTD-----RSAGSLRAMPANLGVTEDEA 177
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 179 AVAVSSNEFPADNPLKFTGSKALSKFHVHVTXLDHGRRRAGKQLRTWMAHDL--BDPA 235
QY 178 TGAATAIRITDLS-----RDLTITOG 198
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 236 TGSASCTLAAYMHIHGADKOTGTQRRAREFTOG 267

```

```

RESULT 6
US-09-738-626-6518
; Sequence 6518, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI

```

```

; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 6518
; LENGTH: 280
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-6518

```

```

Query Match      13.4%; Score 159.5; DB 3; Length 280;
Best Local Similarity 25.9%; Pred. No. 6.4e-08;
Matches 73; Conservative 33; Mismatches 95; Indels 81; Gaps 15;

QY 10 VFPTSDGNFNGNPLGV--NASKVHRDROQLAAGSGSETIFVLDPSBSTTAATITHTP 68
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 14 VFS--SEPFMGNPLAVIADDDLSEQMARLARMTNLSSETFLKPTQEGADYVRIFTP 72
QY 69 TELPAGHPTVGSWMLR-----ERGTPIINTLQVAGIVQ--SYHDDLTAISRSWAP 123
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 73 GELPFGHPTLGAHVFRLEHGQGTOL--VQSCVAGLVAARALDPASGLAFOA--PTL 129
QY 124 AIHDLPSLALAAADPADF--PDDI--AHYLMWTWDRSAG----- 159
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 130 KQGPDLASDLDAACBALGISPDFIRAH---QWVDNGPMVAVELPSAQHYLDLEPPSAH 166
QY 160 -----SLRAMPANLGVTEDEATGAARITDYLSPRLITTOG----- 198
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 187 PTLKLVITGAYPEGADPAFEVRAFOGIG--BDPVTGSLNASTIOWLHRDRAGABGYLAS 244
QY 199 KGSLLI-----HTTWSPEGWVAVAGRY--VSDGVAQ 226
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 245 QGTALGRAGEIHISIESHAIW-----VQGSVTTITFOGTAR 279

```

```

RESULT 7
US-10-389-647-621
; Sequence 621, Application US/10389647
; Publication No. US2004003549A1
; GENERAL INFORMATION:
; APPLICANT: GREENBERG, R. Peter
; APPLICANT: SCHUSTER, Martin
; APPLICANT: LOSTROH, Candi
; TITLE OF INVENTION: QUORUM SENSING SIGNALING IN BACTERIA
; FILE REFERENCE: UIZ-038CP
; CURRENT APPLICATION NUMBER: US/10/389,647
; CURRENT FILING DATE: 2003-03-14
; PRIOR APPLICATION NUMBER: 09/653730
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/153022
; PRIOR FILING DATE: 1999-09-03
; NUMBER OF SEQ ID NOS: 710
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 621
; LENGTH: 278
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa

```

US-10-389-647-621

Query Match 12.7%; Score 151.5; DB 4; Length 278;
Best Local Similarity 28.1%; Pred. No. 4.4e-07;
Matches 62; Conservative 26; Mismatches 88; Indels 45; Gaps 9;

QY 19 GNPVGV-1NASKYEHNDROOLAAOSGSEITFPVDSFGSTTAAHTHTPTTEIPFAGHP 77
DB 17 GNPVAVFPDCCDLSGEMORMAREMNLSESTFVLRPQODG-DARIRIFTFVNLFPAGHP 75
QY 78 TVGASWML-RERGTPINTLOVPAGIVSYH---GDLTAISAS---EW----- 119
DB 76 LIGTALALAGATKDLFETRMGTYPFALERODKVVACSMQPTPTWHSRPAELIA 135
QY 120 -----APEPAIHDLSDLAADP-----ADPDDIAHYLTWTDRSA 158
DB 136 ALGLKSTPPIEVYRNGPRHVPGLSVALSALHPDRALCPFP-DLAVNCFAGAGR-- 192
QY 159 GSILRAMPANLGVTEDEATGAAAIRTTYLSRDLTTTQCK 199
DB 193 -HWRSHMFPSPAYGVVEDAATGSAAGPLAHLAHRROIPLYGQ 232

RESULT 8

US-10-156-761-14521
; Sequence 14521, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 14521
; LENGTH: 272
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-14521

Query Match 10.7%; Score 128; DB 4; Length 272;
Best Local Similarity 23.6%; Pred. No. 0.00013;
Matches 68; Conservative 32; Mismatches 100; Indels 88; Gaps 13;

QY 3 IEYSVLKVTDSGNFNLGVINASKV-EHRDROOLAAOSGSEITFPD-LPSGSTTA 60
DB 1 MRRIRYDAFTDRP-FAGNPAGVLLDPAFPDDAWLQVAREVNNAFAFARLRGSGGSA 59
QY 61 HATIH--TPRTEIPFAGHPVGSWMLBERGT----- 90
DB 60 DWALRFTATREYAMCGHATLATVAHLVHTSGTGEFVRPATRSGLIAPREDGSTITLD 119
QY 91 ---PINTLOVPAGIVSYHGDLTAAISAS---EWAPPAIHDL-DSLDAAL----- 135
DB 120 PPAFLPVEIPREVAALAGAPLITACDTGNVDDLIELADEKTVAGLAPDLRALARYGE 179
QY 136 -----AADPAFPDDIAHYLTWTDSAGSLRAMPANLGVTEDEATGAAAIRITD 187
DB 180 RGIATARAADPAGHDFV-----SRCFEPNIGIDEDPVTGSAHALAP 223
QY 188 YLSRDLTTTQKSLHTTWSF-EGWVR-----VAGRVYS--DG 223
DB 224 FWSQRL-----GRDPLTGLQASPRSGHVRTELDRDRTLTLGRAVTVDG 267

RESULT 9

US-10-425-115-230986
; Sequence 230986, Application US/10425115
; Publication No. US20040214272A1

QY 35 ROOLAAOSGSEITFPVDSFGS-TTAAHTHTPTTEIPFAGHPVGSW---LBERGT 90
DB 4 KOEIAKEFNYSTIFVHPVDPTQORIDITFTREBLEFPAGHPYTGASWALYISPDKG 63
QY 91 PINTLOVPAGIVSYH--GDLTA 112
DB 64 DVTILLTKAGPIPIGLRPSGDEAA 87

US-10-425-115-230986
; OTHER INFORMATION: Clone ID: MRT4577_14252C.1.pcp

Query Match 10.5%; Score 125; DB 4; Length 120;
Best Local Similarity 39.3%; Pred. No. 8.4e-05;
Matches 33; Conservative 11; Mismatches 34; Indels 6; Gaps 3;

QY 35 ROOLAAOSGSEITFPVDSFGS-TTAAHTHTPTTEIPFAGHPVGSW---LBERGT 90
DB 4 KOEIAKEFNYSTIFVHPVDPTQORIDITFTREBLEFPAGHPYTGASWALYISPDKG 63
QY 91 PINTLOVPAGIVSYH--GDLTA 112
DB 64 DVTILLTKAGPIPIGLRPSGDEAA 87

RESULT 10

US-10-166-087-30
; Sequence 30, Application US/10166087
; Publication No. US200300776767A1

QY 35 ROOLAAOSGSEITFPVDSFGS-TTAAHTHTPTTEIPFAGHPVGSW---LBERGT 90
DB 4 KOEIAKEFNYSTIFVHPVDPTQORIDITFTREBLEFPAGHPYTGASWALYISPDKG 63
QY 91 PINTLOVPAGIVSYH--GDLTA 112
DB 64 DVTILLTKAGPIPIGLRPSGDEAA 87

US-10-166-087-30
; OTHER INFORMATION: Genes and proteins for the biosynthesis of anthracyclin

Query Match 9.4%; Score 112.5; DB 4; Length 274;
Best Local Similarity 24.8%; Pred. No. 0.0054;
Matches 52; Conservative 22; Mismatches 71; Indels 65; Gaps 10;

QY 40 AOSGSEITFPD---LPSGSTTAAHTHTPTTEIPFAGHPVGSWMLBERGTINT 94
DB 99 AESG-SQWLMTDGHVRYVFPDVPVHNRGIRHPDPAFGPYAGTPYAGV-----GRAFL 151
QY 95 IQVPAGIVSYHGDLTAAISARSEWAPPAIHDLSDLAADPADPPD---DIAHYL 150
DB 152 LRV-----ADPRTLPAADPGRRKELGFTDLTVFR 181
QY 151 WTWTRSAGSLRAMPANLGVTEDEATGAAAIR-----TDYLSRDLTTTQK 198

Db 182 W---DPRGVLARVPAFGSGIPEDAGCLPAALALGVALLRLAADRISVTYRQYTV-RG 23

QY 199 KGSLLHTTWSPEGW---VRVAGRVVSDGYA 225
| : | | : | | : | |
Db 238 TESVFRCTGSARGGSANVTITGRVMTGGTA 267

RESULT 11
 US-10-733-969A-87
 Sequence 87, Application US/10733969A
 Publication No. US20040219572A1
 GENERAL INFORMATION:
 APPLICANT: CHEN, JIE
 APPLICANT: HU, LIPING
 APPLICANT: LIU, TONG HUA
 APPLICANT: LU, ZHAO HUI
 APPLICANT: SHEN, YAN
 TITLE OF INVENTION: SPECIFIC MARKERS FOR PANCREATIC CANCER
 FILE REFERENCE: 21525
 CURRENT APPLICATION NUMBER: US/10/733,969A
 CURRENT FILING DATE: 2003-12-11
 PRIOR APPLICATION NUMBER: EP 02028058.2
 PRIOR FILING DATE: 2002-12-17
 NUMBER OF SEQ ID NOS: 110
 SOFTWARE: PatentIn version 3.2
 SEQ ID NO 87
 LENGTH: 288
 TYPE: PRP
 ORGANISM: Homo sapiens
 FEATURE:
 OTHER INFORMATION: MAMD binding protein; Accession NO: as of
 US-10-733-969A-87

[illegible]

RESULT 12 763-52819
US-10-450-763-52819
Sequence 52819, Application US/10450763
Publication No. US20050196754A1
GENERAL INFORMATION:
APPLICANT: Hyseq, Inc
TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
FILE REFERENCE: 790CIP3/US
CURRENT APPLICATION NUMBER: US/10/450,763
CURRENT FILING DATE: 2003-06-11
PRIOR APPLICATION NUMBER: PCT/US01/08631
PRIOR FILING DATE: 2001-03-30
PRIOR APPLICATION NUMBER: 09/540,217
PRIOR FILING DATE: 2000-03-31

```

? PRIOR APPLICATION NUMBER: 09/649,167
? PRIOR FILING DATE: 2000-08-23
? NUMBER OF SEQ ID NOS: 60736
? SOFTWARE: CuiSom
? SEQ ID NO: 52819
? LENGTH: 375
? TYPE: PRT
? ORGANISM: Homo sapiens
US-10-450-763-52819

```

Query Match	9.2%	Score 110;	DB 5;	Length 375;
Best Local Similarity	22.8%	Pred. No. 0.015;		
Matches 53; Conservative	22;	Mismatches 67;	Indels 90;	Gaps 10;

[illegible]

```

RESULT 13
US-10-156-761-13432
? Sequence 13432, Application US/10156761
? Publication No. US20030119018A1
? GENERAL INFORMATION:
? APPLICANT: OMURA, SATOSHI
? APPLICANT: IKEDA, HARUO
? APPLICANT: ISHIKAWA, JUN
? APPLICANT: HORIKAWA, HOROSHI
? APPLICANT: SHIBA, TADAYOSHI
? APPLICANT: SAKAKI, TADAYUKI
? APPLICANT: HATTORI, MASAHIRA
? TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
? FILE REFERENCE: 249-262
? CURRENT APPLICATION NUMBER: US/10/156,761
? CURRENT FILING DATE: 2002-05-29
? PRIOR APPLICATION NUMBER: JP 2001-204089
? PRIOR FILING DATE: 2001-05-30
? PRIOR APPLICATION NUMBER: JP 2001-272697
? PRIOR FILING DATE: 2001-08-02
? NUMBER OF SEQ ID NOS: 15109
? SEQ ID NO 13432
? LENGTH: 636
? TYPE: PRT
? ORGANISM: Streptomyces avermitilis
US-10-156-761-13432

```

Query Match	8.4%;	Score 100.5;	DB 4;	Length 636;
Best Local Similarity	24.5%;	Pred. No. 0.31;		
Matches 64;	Conservative 29;	Mismatches 97;	Indels 71;	Gaps 15;

```

QY 21 ELGVINAKVEKRDRLQALAQ-----SGVSEIIPFD--LPSSGSTAATITHTP 67
Db 296 PISVV-ARYPEHLDVPAVAADGRITMSDMWDAGSGMAGFVQSGGLASPGA-----GSP 344
QY 68 RTRIP-FAGH-----PTVGASMTLRERTP----INTLQVPAG-----IVQSVYHG 108
Db 349 VTSVARYAGHLDLPVVGIDSRVYSTMTWQSSGMAAFRGVTLVAPPGSTVAVNSRYSDDL 408
QY 109 DL-----TAISARSEWAPR-FAIHDLDSLDALAADPADP-----DDI 146
Db 409 DLTFTSSDGRITSTMWDARTGAADMFTVSGVAAANGATVTAVAARPHLDVFTVVGIDNR 466

```


GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 14, 2006, 17:03:40 ; Search time 258.049 Seconds
(without alignments)
1133.996 Million cell updates/sec

Title: US-10-620-246-70

Perfect score: 3552

Sequence: 1 MAADYDKLFRPHGEMRAPD.....RKYLELAALSDPFRAGRR 666

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_21:*

1: geneseqp1980a:*

2: geneseqp1990a:*

3: geneseqp2000a:*

4: geneseqp2001a:*

5: geneseqp2002a:*

6: geneseqp2003a:*

7: geneseqp2003ba:*

8: geneseqp2004a:*

9: geneseqp2005a:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3552	100.0	666	2 AAW72911	Aaw72911 Mycobacte
2	3552	100.0	666	2 AAY1928	Aay1928 Amino aci
3	3552	100.0	666	4 AAB19846	Aab19846 Mycobacte
4	3552	100.0	666	5 ABUS5989	Abus5989 M. tuberc
5	3552	100.0	666	6 ABUS7020	Abus7020 Protein e
6	1575	44.3	308	2 AAW72930	Aaw72930 Mycobacte
7	1575	44.3	308	2 AAY1947	Aay1947 Amino aci
8	1550	43.6	586	5 ABUS5359	Abus5359 M. tuberc
9	1550	43.6	586	5 ABUS5716	Abus5716 Protein e
10	462.5	13.0	446	4 ABB70063	Abb70063 Drosophi1
11	420	11.8	697	8 ADX72179	Adx72179 Plant ful
12	416.5	11.7	800	8 ABO58564	Abos58564 Human gen
13	416	11.7	598	6 ABUS7030	Abus7030 Protein e
14	410.5	11.6	386	9 AD247759	Ad247759 Mycobacte
15	410.5	11.6	763	2 AAM31852	Aam31852 Mycobacte
16	408.5	11.5	1216	8 ADX37588	Adx37588 Pancreat1
17	408.5	11.5	1250	6 AAO26231	Aao26231 MDDT rela
18	408.5	11.5	1250	6 ADM87156	Adm87156 Human pro
19	402.5	11.3	708	5 ABB91504	Abb91504 Herbicida
20	398	11.2	440	7 ABM90225	Abm90225 Rice abio
21	397.5	11.2	406	4 ABG37250	Abg37250 Novel hum
22	395	11.1	268	4 ABUS2942	Abus2942 Human mam
23	391	11.0	572	2 AAM31855	Aam31855 Mycobacte
24	383.5	10.8	389	8 ADX93736	Adx93736 Plant ful

25	381.5	10.7	256	4 ABUS2945	Abus2945 Human mam
26	380	10.7	277	4 ABUS2936	Abus2936 Human mam
27	375.5	10.6	261	4 ABUS2943	Abus2943 Human mam
28	375.5	10.6	694	5 ABB92421	Abb92421 Herbicida
29	373.5	10.5	258	4 ABUS2941	Abus2941 Human mam
30	373	10.5	693	7 ADM05164	Adm05164 Human pro
31	371.5	10.5	261	4 ABUS2940	Abus2940 Human mam
32	369.5	10.4	639	8 ADN21309	Adn21309 Bacteri1
33	368.5	10.4	335	4 ABG21919	Abg21919 Novel hum
34	367.5	10.3	574	4 AAG67370	Aag67370 Amino aci
35	367.5	10.3	574	4 AAM52322	Aam52322 WABP homo
36	367.5	10.2	1151	4 ABB61598	Abb61598 Drosophi1
37	364	10.2	330	2 AAM57645	Aam57645 Collagen
38	364	10.2	408	2 AAM07539	Aam07539 Collagen
39	364	10.2	408	4 AAB64007	Aab64007 CLP prote
40	364	10.2	408	4 AAB72737	Aab72737 Repetitiv
41	363.5	10.2	270	4 ABUS2939	Abus2939 Human mam
42	363	10.2	639	8 ADN19622	Adn19622 Bacteri1
43	360.5	10.1	1579	8 ADS97981	Ads97981 Protein f
44	359	10.1	262	4 ABUS2944	Abus2944 Human mam
45	358.5	10.1	693	5 ABP69529	Abp69529 Human pol

ALIGNMENTS

RESULT 1

AAW72911 standard; protein; 666 AA.

XX	AAW72911;	
XX	21-JAN-1999 (first entry)	
XX	Mycobacterium tuberculosis antigen RD1-ORF9B.	
XX	Mycobacterium tuberculosis; antigen; vaccine; immunological; immunogen; infection.	
XX	Mycobacterium tuberculosis.	
OS	Mycobacterium tuberculosis.	
XX	PN W09844119-A1.	
XX	08-OCT-1998.	
XX	01-APR-1998; 98WO-DK000132.	
XX	02-APR-1997; 97DK-00000376.	
XX	18-APR-1997; 97US-0044624P.	
XX	10-NOV-1997; 97DK-00001277.	
XX	05-JAN-1998; 98US-0070488P.	
XX	(STAT-) STATENS SERUM INST.	
XX	Andersen P, Nielsen R, Rosenkrands I, Weldingh K, Rasmussen PB; Oettinger T, Florio W;	
XX	WPI; 1998-542705/46.	
XX	N-PDB; AAV63935.	
XX	New isolated mycobacteria polypeptides and nucleic acids - used for developing products for the diagnosis of or vaccination against mycobacterial infections, particularly tuberculosis.	
XX	Claim 1; Page 183-185; 163pp; English.	
XX	The present sequence represents a Mycobacterium tuberculosis protein. Products from the present invention, which describes protein fragments and nucleic acid fragments derived from M.tuberculosis, can be used in the detection of and prevention of mycobacterial infections. In particular, the proteins and nucleic acids can be used for the diagnosis of or vaccination against tuberculosis caused by M. tuberculosis, M. africanum or M. bovis	

```

XX  Sequence 666 AA;
SQ
Query Match      100.0%; Score 3552; DB 2; Length 666;
Best Local Similarity 100.0%; Pred. No. 1.1e-213;
Matches 666; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  1 MAADYDKLFRPHGEMAPDDMAAQPFPDPSAPPPAPASANLKPXGQTPPTSDLSR 60
DB  1 MAADYDKLFRPHGEMAPDDMAAQPFPDPSAPPPAPASANLKPXGQTPPTSDLSR 60
QY  61 FVSAPPPPPPPPPPTPMPIAAGSPSPSPAASKPPTPMPIAAGSPSPSPAASKPPTPMPIA 120
DB  61 FVSAPPPPPPPPPPTPMPIAAGSPSPSPAASKPPTPMPIAAGSPSPSPAASKPPTPMPIA 120
QY  121 IAGBPAPPKPPTPMPIAGPAPTPTESQLABRPPTPTGAPQPSAPAHVSHGP 180
DB  121 IAGBPAPPKPPTPMPIAGPAPTPTESQLABRPPTPTGAPQPSAPAHVSHGP 180
QY  181 HOPRTAPAPPAKMPGIGBPSPSPSPAAPPTPAPQHSRRARRGHRTDTERNV 240
DB  181 HOPRTAPAPPAKMPGIGBPSPSPSPAAPPTPAPQHSRRARRGHRTDTERNV 240
QY  241 GKVATGPTIOARLRABEASGAQLAPGTBSPAPVLGQPRSTLAPPTPAPTPSPSPSPQR 300
DB  241 GKVATGPTIOARLRABEASGAQLAPGTBSPAPVLGQPRSTLAPPTPAPTPSPSPSPQR 300
QY  301 NSGRABRRVHPDLAAQHAAAPDSITTAATTGRRRRRAAPDLATOKSLRPAKGPVKV 360
DB  301 NSGRABRRVHPDLAAQHAAAPDSITTAATTGRRRRRAAPDLATOKSLRPAKGPVKV 360
QY  361 KYVQPKAKTKPPKVVSSQGRMHVHALTRINLGLSPDEKYLDELHARVRNFRGSYQIA 420
DB  361 KYVQPKAKTKPPKVVSSQGRMHVHALTRINLGLSPDEKYLDELHARVRNFRGSYQIA 420
QY  421 VVGLKGAAGKTTTAAAGSTLAQVRADRIALDADPAGNLADRVGRSGATTADYLAEX 480
DB  421 VVGLKGAAGKTTTAAAGSTLAQVRADRIALDADPAGNLADRVGRSGATTADYLAEX 480
QY  481 ELSHNVDIRAHNSVNVNLEVLPAPEYSSAQRALSDADMFITDPAAPFNVLADCGAG 540
DB  481 ELSHNVDIRAHNSVNVNLEVLPAPEYSSAQRALSDADMFITDPAAPFNVLADCGAG 540
QY  541 FPDPLTRGVLTSTVSGVVVAASVSDGAQAQASVALDWLRNNGYODLASRAQVNIHMPGS 600
DB  541 FPDPLTRGVLTSTVSGVVVAASVSDGAQAQASVALDWLRNNGYODLASRAQVNIHMPGS 600
QY  601 PNVAVKDLVRHFEQVQVQPGRVVVMWDRIIAAGTEISLIDLPIYRKVIELAALSDDF 660
DB  601 PNVAVKDLVRHFEQVQVQPGRVVVMWDRIIAAGTEISLIDLPIYRKVIELAALSDDF 660
QY  661 ERAGR 666
DB  661 ERAGR 666

RESULT 2
AAV21928
ID  AAV21928 standard; protein; 666 AA.
XX
XX  AAV21928;
XX
XX  06-SEP-1999 (first entry)
XX
XX  Amino acid sequence of antigen RD1-ORF9B.
XX
XX  Immunogenic; Mycobacterium tuberculosis; immune response; infection;
XX  tuberculosis; fusion polypeptide; T-cell epitope; ESAT-6; MPT59; TB;
XX  pharmaceutical; vaccination; M. africanum; M. bovis; CFP7A; CFP30A;
XX  CFP7B; CFP23; CFP27; CFP30A; RD1-ORF; CFP10A; CFP16; CFP19; CFP23;
XX  CFP25A; CFP30B; CFP7B.
XX
XX  Mycobacterium tuberculosis.
OS

```

```

XX  WO9924577-A1.
XX
XX  20-MAY-1999.
XX
XX  08-OCT-1998; 98WO-DK000438.
XX
XX  10-NOV-1997; 97DK-00001277.
XX  05-JAN-1998; 98US-0070488P.
XX  01-APR-1998; 98WO-DK000132.
XX
XX  (STAT-) STATENS SERUM INST.
XX
XX  Andersen P, Skjot R;
XX
XX  WPI: 1999-347282/29.
XX  DR  N-PSDB; AAX81042.
XX
XX  New immunogenic fragment of Mycobacterium tuberculosis.
XX
XX  Example 2; Page 202-204; 265bp; English.
XX
XX  The invention describes a substantially pure immunogenic polypeptide
XX  fragment (I) from Mycobacterium tuberculosis that is able to evoke a
XX  protective immune response against infections by mycobacteria belonging
XX  to the tuberculosis complex. The invention provides a (1) fusion
XX  polypeptide comprising at least one polypeptide fragment (I) and at least
XX  one fusion partner; (2) a fusion polypeptide fragment comprising a T-cell
XX  epitope from M. tuberculosis protein ESAT-6, or MPT59 and a second
XX  different amino acid sequence from M. tuberculosis, and/or including a
XX  sequence which protects the first amino acid sequence from in vivo
XX  degradation or post-translational processing; (3) a nucleic acid fragment
XX  that encodes the above polypeptides. The polypeptides and nucleic acid
XX  are useful as pharmaceuticals, for diagnosis of and as antigens for
XX  vaccination against TB caused by Mycobacterium tuberculosis, africanum or
XX  bovis. The polypeptides are also useful for diagnosing ongoing or
XX  previous sensitization in an animal with bacteria belonging to the
XX  tuberculosis complex. The invention also describes the use of CFP7A or
XX  CFP30A or a T-cell epitope of for the induction of a strong immune
XX  response in a mammal; use of CFP7B, CFP19 or MPT59-ESAT6 or a T-cell
XX  epitope of for diagnosis of TB in a mammal by performing a DTH type skin
XX  test; use of CFP27, CFP30A, RD1-ORF2, RD1-ORF3, RD1-ORF5, MPT59-ESAT6,
XX  ESAT6-MPT59, CFP10A, CFP16, CFP19, CFP23, CFP25A, CFP30B, CFP7B or a T-
XX  cell epitope of for the preparation of an immunological composition; and
XX  for the preparation of a subunit vaccine
XX
XX  Sequence 666 AA;
SQ
Query Match      100.0%; Score 3552; DB 2; Length 666;
Best Local Similarity 100.0%; Pred. No. 1.1e-213;
Matches 666; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  1 MAADYDKLFRPHGEMAPDDMAAQPFPDPSAPPPAPASANLKPXGQTPPTSDLSR 60
DB  1 MAADYDKLFRPHGEMAPDDMAAQPFPDPSAPPPAPASANLKPXGQTPPTSDLSR 60
QY  61 FVSAPPPPPPPPPPTPMPIAAGSPSPSPAASKPPTPMPIAAGSPSPSPAASKPPTPMPIA 120
DB  61 FVSAPPPPPPPPPPTPMPIAAGSPSPSPAASKPPTPMPIAAGSPSPSPAASKPPTPMPIA 120
QY  121 IAGBPAPPKPPTPMPIAGPAPTPTESQLABRPPTPTGAPQPSAPAHVSHGP 180
DB  121 IAGBPAPPKPPTPMPIAGPAPTPTESQLABRPPTPTGAPQPSAPAHVSHGP 180
QY  181 HOPRTAPAPPAKMPGIGBPSPSPSPAAPPTPAPQHSRRARRGHRTDTERNV 240
DB  181 HOPRTAPAPPAKMPGIGBPSPSPSPAAPPTPAPQHSRRARRGHRTDTERNV 240
QY  241 GKVATGPTIOARLRABEASGAQLAPGTBSPAPVLGQPRSTLAPPTPAPTPSPSPSPQR 300
DB  241 GKVATGPTIOARLRABEASGAQLAPGTBSPAPVLGQPRSTLAPPTPAPTPSPSPSPQR 300
QY  301 NSGRABRRVHPDLAAQHAAAPDSITTAATTGRRRRRAAPDLATOKSLRPAKGPVKV 360

```


PN WO200274903-A2.
 XX 26-SEP-2002.
 PD
 XX 22-FEB-2002; 2002WO-IB001973.
 PF
 XX 22-FEB-2001; 2001US-0270123P.
 PR
 XX (INSP) INST PASTEUR.
 PA
 XX Cole S;
 PI
 XX WPI; 2002-759885/82.
 DR
 XX
 PT Identifying and selecting genes for survival or virulence of mycobacteria
 PT by a comparative genomic analysis of the sequences of Mycobacterium
 PT tuberculosis and M. leprae.
 PS Claim 17; Page 864-865; 874pp; English.
 PS
 XX This invention relates to a novel method for identifying essential genes
 CC for survival or virulence of mycobacteria species. The method comprises
 CC aligning the genomic sequence of a first mycobacterium species on a
 CC genomic sequence of a second mycobacterium species and selecting a
 CC polynucleotide sequence that is highly conserved in both genomes with no
 CC counterparts in other bacterial genomic sequences and that corresponds to
 CC an essential gene for the survival or virulence of mycobacterium species.
 CC The method of the invention is useful for detecting M. tuberculosis or M.
 CC leprae infection. The method reduces the number of potential new targets
 CC and protective antigens for new drugs and vaccine compositions to treat
 CC and prevent mycobacterial diseases, particularly tuberculosis and
 CC leprosy. The present sequence represents a marker protein from
 CC Mycobacterium tuberculosis and Mycobacterium leprae identified using the
 CC method of the invention
 XX
 XX Sequence 666 AA:
 SQ
 Query Match 100.0%; Score 3552; DB 5; Length 666;
 Best Local Similarity 100.0%; Pred. No. 1,1e-213;
 Matches 666; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MAADYDKLFRPHGMAAPDDMAAQPDPDSASPPAPASANLPKPGQTPPPSDLSR 60
 DB 1 MAADYDKLFRPHGMAAPDDMAAQPDPDSASPPAPASANLPKPGQTPPPSDLSR 60
 QY FVSAPPPEPPPEPTPTMPTIAGSPSPSPASPPPTPMPPIAGPBPAPKPTPPMP 120
 DB 61 FVSAPPPEPPPEPTPTMPTIAGSPSPSPASPPPTPMPPIAGPBPAPKPTPPMP 120
 QY 121 IAGPEBAPKPPPTPMPPIAGPAPTPTESQLAAPRPPTPTGAPQOQSPAPHVSHQ 180
 DB 121 IAGPEBAPKPPPTPMPPIAGPAPTPTESQLAAPRPPTPTGAPQOQSPAPHVSHQ 180
 QY 181 HOPRRTPAPRPAKMPGIPSPSPASPPAPPTPAPVSHRRRRRGRRYTTDERV 240
 DB 181 HOPRRTPAPRPAKMPGIPSPSPASPPAPPTPAPVSHRRRRRGRRYTTDERV 240
 QY 241 GVVATGPSIOARLAREKASGOLAPGTBSPAPLAGPSPSTAPPTAPPTPSPSPSP 300
 DB 241 GVVATGPSIOARLAREKASGOLAPGTBSPAPLAGPSPSTAPPTAPPTPSPSPSP 300
 QY 301 NSGRABERVRHDLAAQHAAAPDSITAAATGGRRRRKRAAPDLATOKSLRPAKPKY 360
 DB 301 NSGRABERVRHDLAAQHAAAPDSITAAATGGRRRRKRAAPDLATOKSLRPAKPKY 360
 QY 361 KYKPKPKKTKPKPKVSGQGRWVHALTRINGLSPDEKYEYDLHARRRNRGSGYQA 420
 DB 361 KYKPKPKKTKPKPKVSGQGRWVHALTRINGLSPDEKYEYDLHARRRNRGSGYQA 420
 QY 421 VVGLGGAAGKTTTAAIGSTLAQVRADRIALADADGAGNLADRVRGSGATTAADYLA 480
 DB 421 VVGLGGAAGKTTTAAIGSTLAQVRADRIALADADGAGNLADRVRGSGATTAADYLA 480

QY 481 ELSHYNDIAPARTSVNAANLEFVPAPEYSSAORALSDADNHPFADPAPSRFTYNVLADCGAG 540
 DB 481 ELSHYNDIAPARTSVNAANLEFVPAPEYSSAORALSDADNHPFADPAPSRFTYNVLADCGAG 540
 QY 541 FPDPLTRGVLTSTVSGVWVAASVSDGAQASVALDMLRNGYQDLASRACVYINHIIMP 600
 DB 541 FPDPLTRGVLTSTVSGVWVAASVSDGAQASVALDMLRNGYQDLASRACVYINHIIMP 600
 QY 601 PNVAVEDLVRFEEQVQVOPGRVVMWDRHIAAGTEISLDLDPYKRYLELAALSDP 660
 DB 601 PNVAVEDLVRFEEQVQVOPGRVVMWDRHIAAGTEISLDLDPYKRYLELAALSDP 660
 QY 661 ERAGR 666
 DB 661 ERAGR 666
 RESULT 5
 ID ABU37020 standard; protein; 666 AA.
 AC ABU37020;
 XX
 DT 19-JUN-2003 (first entry)
 XX
 DE Protein encoded by Prokaryotic essential gene #22547.
 XX
 KM Antisense; prokaryotic essential gene; cell proliferation; drug design.
 XX
 OS Mycobacterium tuberculosis.
 XX
 PN WO20027183-A2.
 PD 03-OCT-2002.
 XX
 XX 21-MAR-2002; 2002WO-US009107.
 PR 21-MAR-2001; 2001US-00815242.
 PR 06-SEP-2001; 2001US-00948993.
 PR 25-OCT-2001; 2001US-0342923P.
 PR 08-FEB-2001; 2002US-00072851.
 PR 06-MAR-2002; 2002US-0362699P.
 XX
 PA (ELIT-) ELITRA PHARM INC.
 XX
 PI Wang L, Zamudio C, Malone C, Haseelbeck R, Ohlsen KU, Zykkind JW;
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 PI
 DR WPI; 2003-029926/02.
 DR N-PSDB; ACA40890.
 XX
 PT New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.
 XX
 PS Claim 25; SEQ ID NO 64944; 1766bp; English.
 XX
 CC The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation; (7) identifying a gene in an operon required for
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an

organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 666 AA;

Query Match 100.0%; Score 3552; DB 6; Length 666;
 Best Local Similarity 100.0%; Pred. No. 1.1e-213;
 Matches 666; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAADYDKLFRPHGMEAPDDMAAOPFPDPSPASPPPASANLPPKNCQTPPTSDLSER 60
 DB 1 MAADYDKLFRPHGMEAPDDMAAOPFPDPSPASPPPASANLPPKNCQTPPTSDLSER 60
 QY FVSAPPPPPPPPPPTPMPIAAGBPSPSPAASKCPPTPMPIAGBPAPKPPPTPM 120
 DB FVSAPPPPPPPPPPTPMPIAAGBPSPSPAASKCPPTPMPIAGBPAPKPPPTPM 120
 QY 121 IAGBPAPKPPPTPMPIAGBPAPPTPTESQALAPAPPTPTGAPQGPSPAPVPSHP 180
 DB 121 IAGBPAPKPPPTPMPIAGBPAPPTPTESQALAPAPPTPTGAPQGPSPAPVPSHP 180
 QY 181 HQRRRTAPAPPMKMPGSPSPASRSASAPAPPTPAQHSRARGRRTDTTNNY 240
 DB 181 HQRRRTAPAPPMKMPGSPSPASRSASAPAPPTPAQHSRARGRRTDTTNNY 240
 QY 181 HQRRRTAPAPPMKMPGSPSPASRSASAPAPPTPAQHSRARGRRTDTTNNY 240
 DB 181 HQRRRTAPAPPMKMPGSPSPASRSASAPAPPTPAQHSRARGRRTDTTNNY 240
 QY 241 GKATGPTSIQARIABEASGAQLAPGTSPAPLPGQPSYLAAPTRAPTPPTSPSPQR 300
 DB 241 GKATGPTSIQARIABEASGAQLAPGTSPAPLPGQPSYLAAPTRAPTPPTSPSPQR 300
 QY 301 NSGRABRRVHPDLAAQHAAAPDSITTAATGGRRKRAAPDLATQKSLRPAKGPVK 360
 DB 301 NSGRABRRVHPDLAAQHAAAPDSITTAATGGRRKRAAPDLATQKSLRPAKGPVK 360
 QY 361 KYRPOKPKATKPKRVVSQGRMRHWALTRINGLSPDEKYEELDHAARVRNPRGSYQIA 420
 DB 361 KYRPOKPKATKPKRVVSQGRMRHWALTRINGLSPDEKYEELDHAARVRNPRGSYQIA 420
 QY 421 VVGLKGAAGKTTTLAALGSTLAQVRADRIALADPAGAGLADRVGQSGATTIADVLAEK 480
 DB 421 VVGLKGAAGKTTTLAALGSTLAQVRADRIALADPAGAGLADRVGQSGATTIADVLAEK 480
 QY 481 ELSHYNDIRATSVNANLEVLPAPEYSSAQRALSDADMFHIADPASRFYNLVLAOCGAG 540
 DB 481 ELSHYNDIRATSVNANLEVLPAPEYSSAQRALSDADMFHIADPASRFYNLVLAOCGAG 540
 QY 541 FPDPLTRGVLTSTVSGVVVAVSVSIDGAQASVALDWMIRNNGYODLASRACVIINHIMP 600
 DB 541 FPDPLTRGVLTSTVSGVVVAVSVSIDGAQASVALDWMIRNNGYODLASRACVIINHIMP 600
 QY 601 PNVAVDLVNHFEOQVOPGRVVVMPMDRHTAAGTEISLDLDPYKRVLELAALSDDP 660
 DB 601 PNVAVDLVNHFEOQVOPGRVVVMPMDRHTAAGTEISLDLDPYKRVLELAALSDDP 660
 QY 661 ERAGRR 666
 DB 661 ERAGRR 666

RESULT 6
 AAW72930

ID AAW72930 standard; protein; 308 AA.

XX AAW72930;

XX 21-JAN-1999 (first entry)

XX Mycobacterium tuberculosis antigen RD1-ORF9A.

XX Mycobacterium tuberculosis; antigen; vaccine; immunological; immunogen;

XX infection.

XX Mycobacterium tuberculosis.

XX NO9844119-A1.

XX 08-OCT-1998.

XX 01-APR-1998; 98WO-DK000132.

XX 02-APR-1997; 97DK-00000376.

XX 18-APR-1997; 97US-0044624P.

XX 10-NOV-1997; 97DK-00001277.

XX 05-JAN-1998; 98US-0070488P.

XX (STAT-) STATENS SERUM INST.

XX Andersen P, Nielsen R, Rosenkrands I, Wellingh K, Rasmussen PB,

XX Oettinger T, Florio W;

XX WPI; 1998-542705/46.

XX N-PSDB; AAV63940.

XX New isolated mycobacteria polypeptides and nucleic acids - used for

XX developing products for the diagnosis of or vaccination against

XX mycobacterial infections, particularly tuberculosis.

XX Claim 1; Page 204-205; 163pp; English.

XX The present sequence represents a Mycobacterium tuberculosis protein.

XX Products from the present invention, which describes protein fragments

XX and nucleic acid fragments derived from *M. tuberculosis*, can be used in

XX the detection of and prevention of mycobacterial infections. In

XX particular, the proteins and nucleic acids can be used for the diagnosis

XX of or vaccination against tuberculosis caused by *M. tuberculosis*, *M.*

XX africanum or *M. bovis*

XX Sequence 308 AA;

XX Query Match 44.3%; Score 1575; DB 2; Length 308;

XX Best Local Similarity 99.7%; Pred. No. 1.8e-90;

XX Matches 307; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 359 VKKVPQKPKATKPKRVVSQGRMRHWALTRINGLSPDEKYEELDHAARVRNPRGSYQ 418

DB 1 MKKVKQKPKATKPKRVVSQGRMRHWALTRINGLSPDEKYEELDHAARVRNPRGSYQ 60

QY 419 IAVVGLKGAAGKTTTLAALGSTLAQVRADRIALADPAGAGLADRVGQSGATTIADVLA 478

DB 61 IAVVGLKGAAGKTTTLAALGSTLAQVRADRIALADPAGAGLADRVGQSGATTIADVLA 120

QY 479 EKELSHYNDIRATSVNANLEVLPAPEYSSAQRALSDADMFHIADPASRFYNLVLAOCG 538

DB 121 EKELSHYNDIRATSVNANLEVLPAPEYSSAQRALSDADMFHIADPASRFYNLVLAOCG 180

QY 539 AGFPDPLTRGVLTSTVSGVVVAVSVSIDGAQASVALDWMIRNNGYODLASRACVIINHIMP 598

DB 181 AGFPDPLTRGVLTSTVSGVVVAVSVSIDGAQASVALDWMIRNNGYODLASRACVIINHIMP 240

QY 599 GEPNVAVDLVNHFEOQVOPGRVVVMPMDRHTAAGTEISLDLDPYKRVLELAALSDDP 658

DB 241 GEPNVAVDLVNHFEOQVOPGRVVVMPMDRHTAAGTEISLDLDPYKRVLELAALSDDP 300

QY 659 DPERAGRR 666

DB 301 DPERAGR 308

RESULT 7
ID AAY21947 standard; protein; 308 AA.

AC AAY21947,
XX
DT 06-SEP-1999 (first entry)

DE Amino acid sequence of antigen RD1-ORP9A.

XX Immunogenic; Mycobacterium tuberculosis; immune response; infection;
KM tuberculosi; fusion polypeptide; T-cell epitope; ESAT-6; MPT59; TB;
KM pharmaceutical; vaccination; M. africanum; M. bovis; CFP7A; CFP30A;
KM CFP7B; CFP19; CFP27; CFP30A; RD1-ORF; CFP10A; CFP16; CFP19; CFP23;
KM CFP25A; CFP30B; CFP7B.

XX Mycobacterium tuberculosis.

OS MO9924577-A1.

PN 20-MAY-1999.

PF 08-OCT-1998; 98WO-DK000438.

PR 10-NOV-1997; 97DK-00001277.

PR 05-JAN-1998; 98US-0070488P.

PR 01-APR-1998; 98WO-DK000132.

PA (STAT-) STATENS SERUM INST.

PI Andersen P, Skjot R;

DR WPI; 1999-347282/29.

DR N-PSDB; AAX81047.

PS Example 2; Page 222-223; 265pp; English.

XX The invention describes a substantially pure immunogenic polypeptide
CC fragment (I) from Mycobacterium tuberculosis that is able to evoke a
CC protective immune response against infections by mycobacteria belonging
CC to the tuberculosis complex. The invention provides a (1) fusion
CC polypeptide comprising at least one polypeptide fragment (I) and at least
CC one fusion partner; (2) a fusion polypeptide fragment comprising a T-cell
CC epitope from M. tuberculosis protein ESAT-6, or MPT59 and a second
CC different amino acid sequence from M. tuberculosis, and/or including a
CC sequence which protects the first amino acid sequence from in vivo
CC degradation or post-translational processing; (3) a nucleic acid fragment
CC that encodes the above polypeptides. The polypeptides and nucleic acid
CC are useful as pharmaceuticals, for diagnosis of and as antigens for
CC vaccination against TB caused by Mycobacterium tuberculosis, africanum or
CC bovis. The polypeptides are also useful for diagnosing ongoing or
CC previous sensitization in an animal with bacteria belonging to the
CC tuberculosis complex. The invention also describes the use of CFP7A or
CC CFP30A or a T-cell epitope of for the induction of a strong immune
CC response in a mammal; use of CFP7B, CFP19 or MPT59-ESAT6 or a T-cell
CC epitope of for diagnosis of TB in a mammal by performing a DTH type skin
CC test; use of CFP27, CFP30A, RD1-ORF2, RD1-ORF3, RD1-ORF5, MPT59-ESAT6,
CC ESAT6-MPT59, CFP10A, CFP16, CFP19, CFP23, CFP25A, CFP30B, CFP7B or a T-
CC cell epitope of for the preparation of an immunological composition; and
CC for the preparation of a subunit vaccine

XX Sequence 308 AA;

Query Match 44.3%; Score 1575; DB 2; Length 308;
Best Local Similarity 99.7%; Pred. No. 1.8e-50;
Matches 307; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 359 VKKVKQKXKATKPPKPVVSQKGRHWVHALTRINLGLSPDEKVELDLHARVRNPRGSGY 418
DB 1 MKKVKQKXKATKPPKPVVSQKGRHWVHALTRINLGLSPDEKVELDLHARVRNPRGSGY 60
QY 419 IAVVGLKGGAGKTTLLAALGSTLLAQRADRIIALADPADAGNLADRVGRSGATTIADVLA 478
DB 61 IAVGLKGGAGKTTLLAALGSTLLAQRADRIIALADPADAGNLADRVGRSGATTIADVLA 120
QY 479 EKELSHYNDIRAHTSVNAVNLEVLPAPEYSSAQALSDADWHEIADPASRPYNLVADCG 538
DB 121 EKELSHYNDIRAHTSVNAVNLEVLPAPEYSSAQALSDADWHEIADPASRPYNLVADCG 180
QY 539 AGFPDPLTGTSTVSGVVVVASVSDGAQOASVALDMTRNNGYODLASARACVIRHIMP 598
DB 181 AGFPDPLTGTSTVSGVVVVASVSDGAQOASVALDMTRNNGYODLASARACVIRHIMP 240
QY 599 GEPNVAVKDLVNHFEQOVQGRVWVWPMRHHIAAGTISLDDLDPYKXKVLLEAALASD 658
DB 241 GEPNVAVKDLVNHFEQOVQGRVWVWPMRHHIAAGTISLDDLDPYKXKVLLEAALASD 300
QY 659 DPERAGR 666
DB 301 DPERAGR 308

RESULT 8
ID ABU05359 standard; protein; 586 AA.

AC ABU05359;

DT 08-APR-2003 (first entry)

DE M. tuberculosis and M. leprae marker protein #10.

XX Mycobacterioses; survival; virulence; protective antigen; vaccine;

KM mycobacterial disease; tuberculosis; leprosy.

OS Mycobacterium tuberculosis.

OS Mycobacterium leprae.

PN WO200274903-A2.

PD 26-SEP-2002.

PF 22-FEB-2002; 2002WO-1B001973.

PR 22-FEB-2001; 2001US-0270123P.

PA (INSP) INST PASTEUR.

PI Cole S;

DR WPI; 2002-759885/82.

PT Identifying and selecting genes for survival or virulence of mycobacteria
PT by a comparative genomic analysis of the sequences of Mycobacterium
PT tuberculosis and M. leprae.

PS Claim 17; Page 172-174; 874pp; English.

XX This invention relates to a novel method for identifying essential genes
CC for survival or virulence of mycobacteria species. The method comprises
CC aligning the genomic sequence of a first mycobacterium species on a
CC genomic sequence of a second mycobacterium species and selecting a
CC polynucleotide sequence that is highly conserved in both genomes with no
CC counterparts in other bacterial genomic sequences and that corresponds to
CC an essential gene for the survival or virulence of mycobacterium species.
CC The method of the invention is useful for detecting M. tuberculosis or M.
CC leprae infection. The method reduces the number of potential new targets
CC and protective antigens for new drugs and vaccine compositions to treat
CC and prevent mycobacterial diseases, particularly tuberculosis and
CC leprosy. The present sequence represents a marker protein from

CC Mycobacterium tuberculosis and Mycobacterium leprae identified using the
 CC method of the invention
 CC
 XX
 SQ Sequence 586 AA;

Query Match 43.6%; Score 1550; DB 5; Length 586;
 Best Local Similarity 49.2%; Pred. No. 1.3e-88;
 Matches 335; Conservative 66; Mismatches 168; Indels 112; Gaps 15;

QY 1 MAADYDGLFRPHGEMAPDDMAAOPFPDPSAPPPAPASANLKPNGQTPPTSDLSER 60
 DB 1 MAADYDGLFRDLDGAYASPDQAAGQLFDDAPLVPPIPTCTTNGEVSAP-MPDSQ 59
 QY 61 FVSAPPPPPPPPTPMPPIAGBPPSEBPASKEPTPMPPIAGBEPAPKPTPMP 120
 DB 60 L-----PNNPPAASKSLPMPPI--GSSVQPPPASSESPAPMVSAP---PRSPASLMP 110
 QY 121 IAGBEPAP-----KPTPMPPIAGPATPESQAPPPPTPTGAPQOP 169
 DB 111 ISBPPQWPAABEPHQPAPAKAPPEVPPIPINEPSP-----AKETMPMTPTDGSOR-- 161
 QY 170 SPAPVHSHGPHORRTAPAPAPMAKMPIGSP--PAPS---RPSAPBPTPAPQHSR 224
 DB 162 TPV-----TSPESLAEFR-AQPPATPKESLPRPMSSPEEPAPPSAQHSR 207
 QY 225 RABRGHRTDTERNVGKVAITPESIQARLAEASGAQLAFTESPAPPLGQPRSYLAP 284
 DB 208 HAARGHHHROETQOANPASATEPPIAPARTAE---LRQAPHAABAP----- 253
 QY 285 TRAPLTPPPSPSPFQNSGRBAERVRHVDLAQAHAAPDSTIATTTGRRRKAPDLD 344
 DB 254 -----TQHLTRPDGLVSHRTALHDS-TATSAIGVQGTGRST---- 287
 QY 345 ATQKSLRAKAPKPVKKVQKPKATKPPKVSQGRHWHVHATRLNLGSPKEXELD 404
 DB 288 -----GAKKPSKVAAKGRHWHVHATRLNLGSPKEXELD 324
 QY 405 LHAVERRNPRGSYOIAVVGKKGAGAKTTLTAAGSTLAQVADRIILADDPGAGNADR 464
 DB 325 LRRVRRRPRGSGYOIGILGKGGAKTIVTIGSMFAARRNDRIIVVADDSGNADR 384
 QY 465 VEGSGATIAVIAKELSHYNDIRATSVANVLEVLPAPEYSSAQRALSDADWHFIAD 524
 DB 385 AGFSEKANIADLADKQVKSYNDIRATSVANVLEVLPAPEYSTAQHALSGEDMNPAAA 444
 QY 525 PARFNVLVADGAGFPDPLTGTGVSTVGCVVVAVSIDGAQAASVALDMLNNGYOD 584
 DB 445 TVSKYVNVMLADCGVGLFDPVTRGVLTASGVVLTISTVDAAQAALADMLNNGYOD 504
 QY 585 IASRACVVNHIMGEENVAVKDLVRFEEQOVQGRVVVWMPMDRIIAAGTEISIDLDP 644
 DB 505 LLSRACVVNHVWKEPNVASKDLVQFEEQIOGRVVVLPMDRIIAAGTEIRLDRDPL 564
 QY 645 YKRKVELAAALSDDFERAGR 665
 DB 565 YRRRIELEAAALSDDFERAGR 585

RESULT 9
 ABU35716
 ID ABU35716 standard; protein; 586 AA.
 AC
 XX ABU35716;
 DT 19-JUN-2003 (first entry)
 XX
 XX Protein encoded by prokaryotic essential gene #21243.
 DB
 XX Antisense; prokaryotic essential gene; cell proliferation; drug design.
 KM
 OS Mycobacterium leprae.
 OS
 PN WO200277183-A2.

XX 03-OCT-2002.
 PD
 XX
 PF 21-MAR-2002; 2002WO-US009107.
 XX
 XX
 PR 21-MAR-2001; 2001US-00815242.
 PR 06-SEP-2001; 2001US-00948993.
 PR 25-OCT-2001; 2001US-0342923P.
 PR 08-FEB-2002; 2002US-00072851.
 PR 06-MAR-2002; 2002US-0362699P.
 XX
 PA (ELIT-) ELITRA PHARM INC.
 XX
 PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KU, Zyskind JW,
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Foreyth RA, Xu HH;
 XX
 DR WPI; 2003-029926/02.
 DR N-PSDB; ACA39586.
 PT New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 XX isolate candidate molecules for rational drug discovery programs.
 PS Claim 25; SEQ ID NO 63640; 1766pp; English.
 XX
 CC The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX

Query Match 43.6%; Score 1550; DB 6; Length 586;
 Best Local Similarity 49.2%; Pred. No. 1.3e-88;
 Matches 335; Conservative 66; Mismatches 168; Indels 112; Gaps 15;

QY 1 MAADYDGLFRPHGEMAPDDMAAOPFPDPSAPPPAPASANLKPNGQTPPTSDLSER 60
 DB 1 MAADYDGLFRDLDGAYASPDQAAGQLFDDAPLVPPIPTCTTNGEVSAP-MPDSQ 59
 QY 61 FVSAPPPPPPPPTPMPPIAGBPPSEBPASKEPTPMPPIAGBEPAPKPTPMP 120
 DB 60 L-----PNNPPAASKSLPMPPI--GSSVQPPPASSESPAPMVSAP---PRSPASLMP 110
 QY 121 IAGBEPAP-----KPTPMPPIAGPATPESQAPPPPTPTGAPQOP 169

CC In addition, the probes are used in identifying and characterizing
CC alternative splicing events, in detecting and characterizing gross
CC alterations in the genomic locus that includes their exon, in assessing
CC smaller genomic alterations, in priming the synthesis of nucleic acids,
CC or in expressing the ORF-encoded peptide. The present sequence is a human
CC single exon probe protein of the invention. Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?docId=20030194704

XX Sequence 800 AA;

Query Match 11.7%; Score 416.5; DB 8; Length 800;

Best Local Similarity 36.9%; Pred. No. 1.3e-17;

Matches 101; Conservative 22; Mismatches 84; Indels 67; Gaps 12;

QY 25 PPTDPSASPPAPASANKPKNQTPPTSDLSERFVSAAPPPPPPPPTPTMPPIA 84
DB 579 PLPLPPS---PPPPPSLPPPPPSLPPPPSP-----SLPPPLSLPPPPSP----- 622
QY 85 GSPSPSPASAKPPTPMPDIAGSPAPPKPPTPMPDIAGSPAPPKPPTPMPDIAGAPPT 144
DB 623 SLPPPPPPPSLPPPPPPPP---SLPPPPPSLPPPPPSLSPSLPLPPPPPSLPPPP 681
QY 145 PTESQIAPPPPTPTGTGAPQQPSAPAPVPSHGHPARRTAPADPMAKMPIGSPAP 204
DB 682 PSLLPLPPP-PPSPPPPSLPPPPPPAPPPPLPP-SPQP-----PSBP-----PSBPSPSP 731
QY 205 SRPSASAPPPPTPAPQHSRRARGRKRTDTRANGKATGSGIOARLAEASAGQLA 264
DB 732 SLPPPPSPSP---PSPSLS-----PPPSLSPS----- 757
QY 265 PGTSPAPLGGPASYIAPPTPAPTPSPSPSP 298
DB 758 ----PSPSPPPPPSLSP---PSPSLSPSPSP 785

RESULT 13

ABU37030
ID ABU37030 standard; protein; 598 AA.

XX ABU37030;

DT 19-JUN-2003 (first entry)

DE Protein encoded by prokaryotic essential gene #22557.

KW Antisense; prokaryotic essential gene; cell proliferation; drug design.

OS Mycobacterium tuberculosis.

XX MO200277183-A2.

PN 03-OCT-2002.

PD 21-MAR-2002; 2002WC-US009107.

PF 21-MAR-2001; 2001US-00815242.

PR 06-SEP-2001; 2001US-00948993.

PR 25-OCT-2001; 2001US-0342823P.

PR 08-FEB-2002; 2002US-0007851.

PR 06-MAR-2002; 2002US-0362699P.

XX (ELIT-) ELITRA PHARM INC.

XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KI, Zykynd JW;

XX PI Wall D, Trawick JD, Carr GD, Yamamoto R, Forsyth RA, Xu HH;

XX MPI; 2003-029926/02.

XX N-PSDB; ACA40900.

XX New antisense nucleic acids, useful for identifying proteins or screening

PT isolate candidate molecules for rational drug discovery programs.

XX Claim 25; SEQ ID NO 64954; 1766bp; English.

XX The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 598 AA;

Query Match 11.7%; Score 416; DB 6; Length 598;

Best Local Similarity 28.0%; Pred. No. 1e-17;

Matches 134; Conservative 22; Mismatches 165; Indels 158; Gaps 19;

QY 9 FPHNGMEAPDDMAAPFPDPSASPPAP-----ASANLP----- 43
DB 50 FAPNPAPAPAPLANSPPVAPAVPAPAPIKETWEBAAMPVPAARNKASPPRPAPV 109
QY 44 -KNGGTPTSDLSERFVSAAPPP-----PPPPPPPP- -MPIAGSPSP- -BP 92
DB 110 SRPNPLPPVPPSPSKAAPVVPVPPVMPDPVPPVPIPDRIAPPDPSPSPAP 169
QY 93 AASKPPTPMPDIAGSPAPPKPPTP-----PMPDIAG----- 123
DB 170 VAPWPPPLPPVPPNNHP- PAPSAPVPGVPLAPLPISGRPVAVGSLIALRICCRVCSGV 228
QY 124 -----PEAPPKP-----PTPMPDIAGAPPTPTSQIAPPPPTPTGTGAP 165
DB 229 LAGALNPSRPSGCPKPPAPAVPACAPVPLPLPLPLPLPLPLPLPLPLPLPLPLPL 288
QY 166 QQPSPAPVPSHGHPARRTAPAPWMAKMPIGSP- -PPAPSPASAPAPPTPAPQHS 223
DB 289 APPAPVPPAPAKNDP-----PPAPAPAPASRPAPFAPAPAPASKSPPTPPAPPP 342
QY 224 RRARGHRYRTDTRNVGKATGPSIOARLAEASAGQIAP-----GTBP- -SPAPLQOP 277
DB 343 -----NNVPPSPPTP-----APPPSGDLPPLPAPAPAP 373
QY 278 RSYL-----APPTRAPPTSPSPSPSPQNSGRRAARRVHPDIAQH 318
DB 374 RLSPASPPAPAPFPPTLIMLVPLPPVPPAPAPNSPEPSPAPAPPKMPNPPGPVAPAN 433
QY 319 AAAQPSITAAATTGGRARRAPDIDATQKSAPAAKGVKVKVQKQKATK- -PP 373
DB 434 SPPFPDPPAPVPA- -SVAPAPPTP- -PSANSPFPAPAPAPVAVAPAAANPP 484


```

RESULT 14
ID AD247759 standard; protein: 386 AA.
XX
AC AD247759;
XX
DT 14-JUL-2005 (first entry)
XX
DB Mycobacterium tuberculosis antigen C17 mature protein.
XX
KM anticubercular; tuberculostatic; cloning; immune stimulation; diagnosis;
XX microorganism detection; tuberculosis.
XX
OS Mycobacterium tuberculosis.
XX
PN US6887481-B1.
XX
PD 03-MAY-2005.
XX
PF 15-DEC-1999; 99US-00461774.
XX
PR 16-DEC-1998; 98US-0112499P.
XX
PA (SCTE-) AGENCY SCI TECHNOLOGY & RES.
XX
PI Chan L, Chung MCM, Lim RH;
XX
DR WPI; 2005-312433/32.
XX N-PSDB; AD247758.
XX
PT New isolated polypeptides derived from Mycobacterium tuberculosis and
PT related organisms, useful for diagnosing, preventing or treating
PT tuberculosis or other diseases associated with M. tuberculosis or related
PT organisms.
XX
PS Claim 1; SEQ ID NO 10; 37P; English.
XX
CC The invention relates to an isolated polypeptide obtainable from a
CC species of Mycobacterium and is immunoreactive with sera from a human,
CC animal or avian species previously exposed to the Mycobacterium but is
CC not immunoreactive with sera from a human, animal or avian species not
CC previously exposed to Mycobacterium. The composition and methods are
CC useful for diagnosing, preventing or treating tuberculosis or other
CC diseases or conditions associated with M. tuberculosis or related
CC organisms. This sequence represents the Mycobacterial antigen C17 mature
CC protein.
XX
SQ Sequence 386 AA;

Query Match 11.6%; Score 410.5; DB 9; Length 386;
Best Local Similarity 33.3%; Pred. No. 1.4e-17;
Matches 109; Conservative 20; Mismatches 119; Indels 79; Gaps 12;

QY 15 MEAPD-----DMAQPPFDPSAPPAASANLPKNGQTPTSDLSBRVSA 64
DB 66 LKAPNPVPVPAAPGPNASAPAPMPPTPLPPVPPGSCAPRPVPAVP-----M 113
QY 65 P-----PPPPPPPPPT-----PMDIAGBPPSPPPASKPTPPMPIAGPB-PA 110
DB 114 PPAAPKMWALPPAPAPSPPTSLAVFPVPVPPVPLPYKMPSPVPVPPPPABPTPN 173
QY 111 PPKPPPTPPMIA-GPBAPAPKPPPPM-----PIAGPAPTPTSQLAPRRPTPTPTG 163
DB 174 PPAAPPPPLKNSPPPPVPPVPPVPLTLNPPVPAPPAANTSNPLPAPAPAPLXPG 233
QY 164 APQOBSAPAPVSHGHQRRRTAPAPWAKMPGSEPPAPSPRSASAPBPPTAPADQS 223
DB 234 PPAAPPPAPAPNSAPAP-----SPSPVPVFPPTPPPPPPPPPPNSPPAPAPAPAP-- 286
QY 224 RRAARGHRYTDTERNVGVATGFSIGA-----RLRABEAS--GAQLAPGTSPAPVIGOP 277

```

```

DB 287 -----LPGPSPPAPPAAPPLPNSPAAAPGPPAWGAPDPAP---P 323
QY 278 RSYIAPPTRPAPTPPPSPQPNRSGR 304
DB 324 LPYSSPPAPACVPGAPLAPLPISGR 350

RESULT 15
ID AAW31852 standard; protein: 763 AA.
XX
AC AAW31852;
XX
DT 25-MAR-2003 (revised)
XX
DT 27-APR-1998 (first entry)
XX
DB Mycobacterium tuberculosis 74 kDa protein.
XX
KM Tuberculosis; mycobacteria; infection; diagnosis; antimycobacterial;
XX antibiotic; vaccine.
XX
OS Mycobacterium tuberculosis.
XX
PN W09741252-A2.
XX
PD 06-NOV-1997.
XX
PF 18-APR-1997; 97WO-EP001973.
XX
PR 29-APR-1996; 96DE-01017184.
XX
PA (GBFB ) GBF GES BIOTECH FORSCHUNG GMBH.
XX
PI Singh M, Honisch C, Espitia C, Moreno C;
XX
DR WPI; 1997-549750/50.
XX N-PSDB; AAT93610.
XX
PT New DNA and related proteins or RNA derived from M. tuberculosis - used
PT for diagnosis of mycobacterial infections, monitoring vaccination and
PT development of anti-mycobacterial agents.
XX
PS Claim 5; Fig 13; 55P; English.
XX
CC This novel 74 kDa protein is encoded by an open reading frame of a
CC Mycobacterium tuberculosis DNA fragment (see AAT93610) containing
CC polymorphic GC-rich sequences. Its amino acid sequence shows a high
CC proline content, but there is no homology to any known proline-rich
CC antigens of mycobacteria. Novel M. tuberculosis proteins (see AAW31851-
CC 57) are claimed. These can be produced as recombinant proteins,
CC especially in bacterial, yeast, fungal or higher eukaryote host cells,
CC and used for diagnosing tuberculosis and other mycobacterial infections
CC in humans or animals. The claimed proteins can also be used for
CC epidemiological studies, for monitoring vaccination, and for the
CC development of vaccines and anti-mycobacterial drugs. (Updated on 25-MAR-
CC 2003 to correct PR field.)
XX
SQ Sequence 763 AA;

Query Match 11.6%; Score 410.5; DB 2; Length 763;
Best Local Similarity 29.2%; Pred. No. 3e-17;
Matches 138; Conservative 26; Mismatches 168; Indels 141; Gaps 22;

QY 17 APDDMAAQP-----FDPSAPPPAPASANLP--KXNGCTPTPTSDLSRFVSAPP 69
DB 119 APNSPPPPPPAPAPKVPAPVPVNSPPPPPPAANP-----APAPAP 166
QY 70 -----PPPPPPPP--MPIAGBPPSPPPASKPTPPMPIAGBPA-----PPKPTPP 118
DB 167 LANSPPUPPAPPTPAGTPPAAPVPPVPAAPKSGAPAPRRPAPAPMATPMEFFPLPPVPP 226
QY 119 MPIA-GPBAPAPKPTPPMPIAGP--APTPTSQLAPRRPTPTPTG----- 164

```


GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 14, 2006, 17:19:08 ; Search time 39.0528 Seconds
(without alignments)
1640.866 Million cell updates/sec

Title: US-10-620-246-70

Sequence: 1 MAADYDKLFRPHEGMBAPDD.....RKVLLELAALSDDFERAGR 666

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Minimum Match 0%

```
Database :
1:  PIR_80:*
2:  p1r1:*
3:  p1r2:*
4:  p1r3:*
5:  p1r4:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	3552	100.0	666	2	B70803	hypothetical protee
2	1550	43.6	586	2	H86314	conserved hypothet
3	1378.5	38.8	458	2	T10030	hypothetical prote
4	530	14.9	405	2	B70545	hypothetical prote
5	421	11.9	1188	2	S49315	extensin-like prote
6	414.5	11.7	1201	2	G86441	unknown protein f
7	413	11.7	760	2	T06291	extensin homolog
8	414	11.6	444	2	S26597	extensin - Volvox
9	402.5	11.3	708	2	D96718	hypothetical prote
10	391.5	11.0	582	2	G87008	conserved hypothet
11	389	11.0	907	2	B96636	hypothetical prote
12	386.5	10.9	341	2	A70598	hypothetical prote
13	383	10.8	330	2	A70656	hypothetical prote
14	381.5	10.7	760	2	F86387	probable Pro kina
15	377	10.6	839	2	F75518	hypothetical prote
16	375	10.6	630	2	S06733	hydroxyproline-ric
17	371	10.4	317	2	A28996	proline-rich prote
18	371	10.4	350	2	S22456	hydroxyproline-ric
19	368.5	10.4	376	2	F71558	probable cell wall
20	367.5	10.3	574	2	T43556	Wiskott-Aldrich sy
21	364.5	10.3	574	2	T38819	wiskott-aldrich sy
22	364	10.2	544	2	T17547	proline-rich prote
23	357	10.1	267	2	S08314	cell wall glycopro
24	355.5	10.0	1585	2	T31611	hypothetical prote
25	354.5	10.0	489	2	T11622	extensin class 1 p
26	353.5	10.0	280	2	T11671	extensin-like prot
27	353	9.9	744	2	B86255	hypothetical prote
28	349	9.8	847	2	F96531	hypothetical prote
29	348	9.8	567	2	B70884	hypothetical prote

ALIGNMENTS

30	343.5	9.7	727	2	CA4534	hypotheical prote
31	343	9.7	416	2	JU0465	extensin precursor
32	343	9.7	710	2	D56728	hypotheical prote
33	342	9.6	240	2	B24264	proline-rich prote
34	341	9.6	240	2	A24264	proline-rich prote
35	340.5	9.6	283	2	S13383	hydroxyproline-ric
36	340	9.6	839	2	T04859	extensin homolog F
37	339.5	9.6	353	2	S36438	BPPT protein - hyd
38	338.5	9.5	461	2	T10741	extensin-like prot
39	338	9.5	303	2	S28264	hydroxyproline-ric
40	337.5	9.5	328	2	J00985	hydroxyproline-ric
41	337.5	9.5	1611	2	T82826	hypotheical prote
42	337	9.5	895	2	C63711	99.7K hypotheical
43	336.5	9.5	761	2	CA4672	hypotheical prote
44	335.5	9.4	369	2	S20500	hydroxyproline-ric
45	334.5	9.4	539	2	T28770	hypotheical prote

RESULT 1

hypothetical protein RV3876 - *Mycobacterium tuberculosis* (strain H37Rv)
C:Species: *Mycobacterium tuberculosis*
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C:Accession: B70803
R: Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holtzoyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S. Nature 393, 537-544, 1998
A: Authors: Scares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A: Title: Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome
A: Reference number: A70500; MUID:98295987; PMID:9634230
A: Accession: B70803
A: Status: preliminary; nucleic acid sequence not shown; translation not shown
A: Molecule type: DNA
A: Residues: 1-666 <COL>
A: Cross-references: UNIPROT:O69740; UNIPARC:UPI00000D105D; GB:AL022120; GB:AL123456; NID
A: Experimental source: strain H37Rv
C: Genetics:
A: Gene: RV3876

Query Match	100.0%;	Score 3552;	DB 2;	Length 666;
Best local Similarity	100.0%;	Pred. No. 3	6e-153;	
Matches 666;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0
QY	1	MAADYDKLFRHBEKAPDDMAAQPFDPSSAPPASANLKPNGQTPPTSDLSR	60	
Db	1	MAADYDKLFRHBEKAPDDMAAQPFDPSSAPPASANLKPNGQTPPTSDLSR	60	
QY	61	FVSAPPPPPPPPPPPPTPWFIAAGBPSSBPAAKSPPTPMPFIAGBEPAPKPTPMP	120	
Db	61	FVSAPPPPPPPPPPPPTPWFIAAGBPSSBPAAKSPPTPMPFIAGBEPAPKPTPMP	120	
QY	121	IAGBPAPKPTPMPPIAGAPPTPTSSQLAPRRPPPTQPTTGAPOQSSPAHVSHRP	180	
Db	121	IAGBPAPKPTPMPPIAGAPPTPTSSQLAPRRPPPTQPTTGAPOQSSPAHVSHRP	180	
QY	181	HOPRRTAAPWAMPPIGEBPPAPSRSSAPABPPTPAPOHSRRARRGRRTDTERAV	240	
Db	181	HOPRRTAAPWAMPPIGEBPPAPSRSSAPABPPTPAPOHSRRARRGRRTDTERAV	240	
QY	241	GKVATGPSIQARLARBESAGQLAPGTSPAPLPGQPSYLAIPTRPAPTSPSSPOR	300	
Db	241	GKVATGPSIQARLARBESAGQLAPGTSPAPLPGQPSYLAIPTRPAPTSPSSPOR	300	
QY	301	NSGRABRRVAPDIAAQAAPDSITPAITGGRRRKRAAPDDIATQSLRPAKGPVK	360	
Db	301	NSGRABRRVAPDIAAQAAPDSITPAITGGRRRKRAAPDDIATQSLRPAKGPVK	360	
QY	361	KYKPOKPAATPEPKVNSQGRHWHVALTINTGLSPDEKTELDLHAYRBNRGSYQIA	420	
Db	361	KYKPOKPAATPEPKVNSQGRHWHVALTINTGLSPDEKTELDLHAYRBNRGSYQIA	420	

```
Db 361 KKKPKKATKPKPKVVSQKGRHWVHALTRINIGLSPDEKTEYLDLHARVRNRGSGYQIA 420
Qy 421 VVGLKAGAKTTLTAALGSTLAQVRADRIALADPADGAGNLADRVKRGSGATTADYLAEK 480
Db 421 VVGLKAGAKTTLTAALGSTLAQVRADRIALADPADGAGNLADRVKRGSGATTADYLAEK 480
Qy 481 ELISHNDIDRAHTSVNAVLEVLPAPEYSSAQRLSDADMHFIADPASFNNVLADCCGAG 540
Db 481 ELISHNDIDRAHTSVNAVLEVLPAPEYSSAQRLSDADMHFIADPASFNNVLADCCGAG 540
Qy 541 FPDPLTRGLSTVSGVVAASVSDGAQOASVALDMLRNNGYODLASRACVVIINHIMPG 600
Db 541 FPDPLTRGLSTVSGVVAASVSDGAQOASVALDMLRNNGYODLASRACVVIINHIMPG 600
Qy 601 PNVAVDLYRHFEOQVOPGRVVMVPMWRHIAAGTEISLIDLPIYRKRYELAAALSDPF 660
Db 601 PNVAVDLYRHFEOQVOPGRVVMVPMWRHIAAGTEISLIDLPIYRKRYELAAALSDPF 660
Qy 661 ERAGRR 666
Db 661 ERAGRR 666
```

RESULT 2

```
H86914
conserved hypothetical protein ML0048 [imported] - Mycobacterium leprae
C:Species: Mycobacterium leprae
C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #ext_change 09-Jul-2004
C:Accession: H86914
R:Coile, S.T.; Eiglmeyer, K.; Parkhill, J.; James, K.D.; Thomson, N.R.; Wheeler, P.R.; Ho
R.; Davies, R.M.; Devlin, K.; Duthey, S.; Felwell, T.; Fraser, A.; Hamlin, N.; Holtoy,
eam, M.A.; Rutherford, K.M., 2001
Nature 409, 1007-1011, 2001
A:Author: Rutter, S.; Seeger, K.; Simon, S.; Simmonds, M.; Skelton, J.; Squares, R.; Sq
A:Title: Massive gene decay in the leprosy bacillus.
A:Reference number: A86909; PMID:21128732; PMID:11234002
A:Accession: H86914
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-566 <STO>
A:Cross-references: UNIPROT:Q9CDB8; UNIPARC:UPI000006C623; GB:AL450380; NID:913092445; F
C:Genetic:
A:Gene: ML0048
```

```
Query Match 43.6%; Score 1550; DB 2; Length 586;
Best Local Similarity 49.2%; Pred. No. 4.5e-63;
Matches 335; Conservative 66; Mismatches 168; Indels 112; Gaps 15;

Qy 1 MAADYDKLFRPHGMEAPDDMAAQPDPSPASFPAPASANLPEKNGQTPPTSDLSER 60
Db 1 MAADYDKLFRLDGAVASPDQAEOLEFDAPLYPPIIPTCTTPNGEVAS-P-MPDWSEQ 59
Qy 61 FVSAPRRPPRRPPPTPTMRIAAGBPSBPASASRPTTPMRIAQBEPAPRRPPRPM 120
Db 61 FVSAPRRPPRRPPPTPTMRIAAGBPSBPASASRPTTPMRIAQBEPAPRRPPRPM 120
Qy 60 L-----PBNPPASKSPLPMPPI--GSSVOPPPASSSPAPMPVASP--PRSPASLMP 110
Db 60 L-----PBNPPASKSPLPMPPI--GSSVOPPPASSSPAPMPVASP--PRSPASLMP 110
Qy 121 IAGBEPAP-----KPTTPMRIAQBEPAPTESQADPRRPTPTPTGAPQOE 169
Db 121 IAGBEPAP-----KPTTPMRIAQBEPAPTESQADPRRPTPTPTGAPQOE 169
Qy 111 ISEPPMPAPABENHOPAKAEPSPVPIPIHPSR-----ACAPMTMTPTIDSGQR-- 161
Db 111 ISEPPMPAPABENHOPAKAEPSPVPIPIHPSR-----ACAPMTMTPTIDSGQR-- 161
Qy 170 SPAPVHSHGPHQPRRTAPAPMAKMPIGEP--PAPS--RPSAPBEPPTPAPQHSR 224
Db 170 SPAPVHSHGPHQPRRTAPAPMAKMPIGEP--PAPS--RPSAPBEPPTPAPQHSR 224
Qy 162 TPV-----TSPPSLAEPF-AQPPATPKSLPRPMSAPPEAPRPSANQHSR 207
Db 162 TPV-----TSPPSLAEPF-AQPPATPKSLPRPMSAPPEAPRPSANQHSR 207
Qy 225 PARGRHRTDTERAVNGKATGTSIQARLAEASGAQLAQTGEBPAPRIGQRSYLAEP 264
Db 225 PARGRHRTDTERAVNGKATGTSIQARLAEASGAQLAQTGEBPAPRIGQRSYLAEP 264
Qy 208 HARRGHHDDETQOANPASATEPMIAPRATAE--LRQAPHAABAP----- 253
Db 208 HARRGHHDDETQOANPASATEPMIAPRATAE--LRQAPHAABAP----- 253
Qy 285 TRBAPTEPPSPSPQNSGRABERAVPDLAQHAQAOPSTIATATGRRKRAAPDLD 344
Db 285 TRBAPTEPPSPSPQNSGRABERAVPDLAQHAQAOPSTIATATGRRKRAAPDLD 344
Qy 254 -----TQHLTRPDGLVSHRTALHDS-TATSAIGVGTGRST----- 287
Db 254 -----TQHLTRPDGLVSHRTALHDS-TATSAIGVGTGRST----- 287
Qy 345 ATQKSLRPAKGPVKVKKPKATPKPVVSQKGRHWVHALTRINIGLSPDEKTEYLD 404
Db 345 ATQKSLRPAKGPVKVKKPKATPKPVVSQKGRHWVHALTRINIGLSPDEKTEYLD 404
```

```
Db 288 -----GAKKPSKVAKGRHWVHALTRINIGLSPDEKTEYLD 324
Qy 405 LHAVERNRGSGYQIAVVGLKAGAKTTLTAALGSTLAQVRADRIALADPADGAGNLADR 464
Db 325 LHTRVRRPPRGSGYQIGILGLKAGAKTTLVTLGSHFAVRADRIALVADTSCGTLADR 384
Qy 465 VGRSGATTADYLAEKELSHNDIDRAHTSVNAVLEVLPAPEYSSAQRLSDADMHFIAD 524
Db 385 AGRFSEANITADLADDOVSGYNDIRHTSVNAVLEVLPAPEYSTQRLSGEDNMFALA 444
Qy 525 PASRFNNVLADCCGAGFPDPLTRGLSTVSGVVAASVSDGAQOASVALDMLRNNGYOD 584
Db 445 TVSKTYNWLADCCGAGLFPDPLTRGLSTASGVVITVSTDADARQALALDMLRNNGYOD 504
Qy 585 LASRACVVIINHIMPGSPNVAVDLYRHFEOQVOPGRVVMVPMWRHIAAGTEISLIDLPI 644
Db 505 LASRACVVIINHIMPGSPNPASLDLYVQFPDQIQOPGVVVVLPMDKHIAAGTEIRLIDLPI 564
Qy 645 YRKRYELAAALSDPFERAGR 665
Db 565 YRRRIELAAALSDPFERAGR 585
```

RESULT 3

```
T10030
hypothetical protein MLCB628.11c - Mycobacterium leprae
C:Species: Mycobacterium leprae
C:Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #ext_change 09-Jul-2004
C:Accession: T10030
R:Eiglmeyer, K.; Honore, N.; Woods, S.A.; Caudron, B.; Cole, S.T.
Mol. Microbiol. 7, 197-206, 1993
A:Title: Use of an ordered cosmid library to deduce the genomic organization of Mycobac
A:Reference number: Z16917; PMID:93188700; PMID:8446027
A:Accession: T10030
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-478 <SIG>
A:Cross-references: UNIPROT:Q33082; UNIPARC:UPI0000086649; EMBL:Y14967; NID:92370268; P
C:Genetic:
A:Note: MLCB628.11c
```

```
Query Match 38.8%; Score 1378.5; DB 2; Length 478;
Best Local Similarity 51.1%; Pred. No. 1.9e-55;
Matches 293; Conservative 53; Mismatches 126; Indels 101; Gaps 11;

Qy 95 SKRPPTPMPRIAQBEP--APKPTTPMRIAQBEPAPKPTTPMRIAQBEPPTESQALAP 152
Db 4 SEPPMP-PALAPENHOPAKAEPSPVPIPIHPSRPAKATPMPTPTIDSGQRTVTS---- 58
Qy 153 PRPPTPTGAPQOEPSPAPVHSHGPHQPRRTAPAPMAKMPIGEPAPRPSASPA 212
Db 59 -----PBNPPASKSPLPMPPI--GSSVOPPPASSSPAPMPVASP--PRSPASLMP 95
Qy 213 EPPTPAPPHSRARHRTDTERAVNGKATGTSIQARLAEASGAQLAQTGEBPAPR 272
Db 213 EPPTPAPPHSRARHRTDTERAVNGKATGTSIQARLAEASGAQLAQTGEBPAPR 272
Qy 96 -----CHSHARRGHHDDETQOANPASATEPMIAPRATAE--LRQAPHAABAP 144
Db 96 -----CHSHARRGHHDDETQOANPASATEPMIAPRATAE--LRQAPHAABAP 144
Qy 273 PLGQPSYLAEPTRAPAPTEPPSPQNSGRABERAVPDLAQHAQAOPSTIATATG 332
Db 145 P-----TQHLTRPDGLVSHRTALHDS-TATSAIGVGTGRST----- 171
Qy 333 GRRRKAAPDLDATOKSLRPAKGPVKVKKPKATPKPVVSQKGRHWVHALTRIN 392
Db 172 GYGTGRST-----GAKKPSKVAKGRHWVHALTRIN 204
Qy 393 LGLSPDEKTEYLDLHARVRNRGSGYQIAVVGLKAGAKTTLTAALGSTLAQVRADRIAL 452
Db 205 LGLSPDEKTEYLDLHARVRNRGSGYQIGILGLKAGAKTTLVTLGSHFAVRADRIAL 264
Qy 453 DADPGAGNLADYVGRSGATTADYLAEKELSHNDIRATSVNAVLEVLPAPEYSSAOR 512
Db 265 DADTSCGTLADRGRFSEANITADLADDOVSGYNDIRHTSVNAVLEVLPAPEYSTQAR 324
```


A:Reference number: AB6141; MID:21016719; PMID:11130712
 A:Accession: G86441
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-1201 <STO>
 A:Cross-references: UNIPROT:Q9C661; UNIPARC:UPI000009C9CB; GB:AE005172; NID:G11136725; F
 C:Genetics:
 A:Map position: 1

Query Match 11.7%; Score 414.5; DB 2; Length 1201;
 Best Local Similarity 24.6%; Pred. No. 1,2e-11;
 Matches 170; Conservative 67; Mismatches 230; Indels 223; Gaps 32;

QY 12 HEGMEAPDDMA-----AQPFDPSASFPAPASAN-----LPRNGQTPPTSDDIS 58
 DB 438 HESSEBPHRSHNNHHIETPAKQSVDMPLNPSDPSSGDHVTLLPRLPRLPRLTSTT 497
 QY 59 ERFVSAAPRRPPPP-----PPRRPTMTI-----AAGEPPSPRPAAS--- 95
 DB 498 SFSPSQPPRRPPPPPLFMSTSPSQPPRRPPPLFTSTSPSQPPRRPPPLPSFSN 557
 QY 96 -----KPPTPMP---IAGP---BPAPKPPTPMPPIAG-----BPAP 128
 DB 558 RPPPLTLHPRIKTPRRPPPPPLPSKSLPPLAOPRRPPPPPPSSSKSTSPSA 617
 QY 129 PKPPTPMPPIAG-----PAPPTP---SOLAPRRPTPOTP-----TGAP 165
 DB 618 PRRPPPPPPSPSGTGKQQAQPPRRPPPPPTPIPAKAKAPRRPPPPPTSHSGSIVGPP 677
 QY 166 QOPESAPAPVP---SHGPHQPKRTAPAP-PAKAMPGEPPPPAPSP-SASAPAPPTPRA 219
 DB 678 STPPPPPPPPPKANISNAPKP---APPLPSPSTRIAGAPPPPPPLSKTPAPPP--- 721
 QY 220 PQHSRRARBGHRYRTDTERVGVKATGSPISQALRAEASGAQLAGTEPSAPLQOPRS 279
 DB 732 -----PLSKTPVPPPPPG-----LKGSTSGPPPLAKGS 761
 QY 280 YLAPTRPAPTEPPSPSPQNSGRRAERKV-PDLAQHAAAP---DSITATTGG--- 333
 DB 762 NAPP-----PPRAGRGASLGLGRGVSVPTAAPKTALKPLHMSKVTRAAKSLW 814
 QY 334 -----RRRKRAADLDATQ-KSLRPAKGPVKVAPQKPKA-TKPPKV---VSGRGKRW 384
 DB 815 AUTQKQENQPRAPEDISLESLSFSAVSDTTAKSGTGRSSLSKKEKQVLVDLRRANNC 874
 QY 385 VHALTRINGLASPD-----SKVEL-----DLHAR 408
 DB 875 EIMLTIKIKIPL-DMMLSAVLALSLALDIDQENLIKFCPTKEMELLRYTGDKEMLCK 933
 QY 409 VRNPPRGSYQIAVVGKKGAGK-----TTTLAAGSTTLAQVRADRII-----A 451
 DB 934 CEVPPRIEAKLRFGFKITPASQVBEKSGCLNTINATKEKVSAKLRQIMQITILTGNA 993
 QY 452 LDADPAGN-----LADRVGSGATTADVLAE-----KELSHYNDIRAHYS 493
 DB 994 LINGSTARGSAGVFKLDSLKLSDTRANNKMTILMHYLCIKVGEKMBELDFANDVLHLEA 1053
 QY 494 VNAVNLVLPAPYSSAQRLSDADMHFIA 523
 DB 1054 ASKIELEKTL-AEEMQATKGLKEVQEELMA 1082

RESULT 7

T06291
 exensin homolog T9E8.80 - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 09-Jul-2004
 C:Accession: T06291
 R:Bevan, M.; Pohl, T.; Weisenegger, T.; Bannroff, I.; Mewes, H.W.; Mayer, K.F.X.; Schuel
 submitted to the Protein Sequence Database, April 1999
 A:Reference number: Z15588
 A:Accession: T06291
 A:Molecule type: DNA

A:Residues: 1-760 <BV>
 A:Cross-references: UNIPROT:Q9T0K5; UNIPARC:UPI000009FB81; EMBL:AL049608
 A:Experimental source: Cultivar Columbia; BAC clone T9E8
 C:Genetics:
 A:Map position: 4
 A:Note: T9E8.80

Query Match 11.7%; Score 414; DB 2; Length 760;
 Best Local Similarity 29.9%; Pred. No. 8,2e-12;
 Matches 123; Conservative 20; Mismatches 145; Indels 124; Gaps 18;

QY 17 ABDDMAAPFPDPSAFPPAPASANLPKNGQTPPTSDLSBRFASAPP---PPPPP 73
 DB 393 SPAPVVTLLPPSLSPPPPPAPISFTPTLSSPP-----SPRPVYSPPPPP 443
 QY 74 PRRPTPPIAGSPSPSEBPAASKPPTPMPPIAGP---BPAPKPPTPMPPIAGPEBAPK 130
 DB 444 PRRPVYSPRRPPPPPPPVYSPRRPPPPPPPVYSPSPSPRRPPPPPPPVYSPRRPPPP 503
 QY 131 PRRPPM-----PIAGPATPTESQLAPRRP-TTOTPGARQOPE----- 169
 DB 504 PRRPVYSPRRPVYSPRRPPSPAPTPVYCTRRPPPPSHSPPRPFSPRRPVYSSAP 563
 QY 170 -----SPAPH---VPSHGPPQPRRTAPAPPAKMPGEPAPR--SRPSAPR-EPPTPRA 219
 DB 564 PRRHSPRRHSPRRPPPPPPPVYSPRRPTPVSSPPPTPVYSPRRPPCTIBPPPP- 622
 QY 220 PQHSRRARBGHRYRTDTERVGVKATGSPISQALRAEASGAQLAGTEP-----SPAP 273
 DB 623 -----PCI-----EYSPPPPPPVHYSPPP 643
 QY 274 LQOPNSYLAPPTPRA-PTPPSPSPQNSGRRAERKV-PDLAQHAAAPDSITATT 331
 DB 644 --PPVYSSPPPPPVYSSPPPPPVYSSPPPEVHYNSPPSPPVYSSPPPPSA--- 698
 QY 332 GGRRRRAAPDLDATQSLRPAKGPVKV---KPQKPKA--TKPPKVSQ 378
 DB 699 -----PCESPPRAPVNHSPPPPMHNSPPPPVTHQ 730

RESULT 8

S22697
 exensin - Volvox carteri (fragment)
 C:Species: Volvox carteri
 C:Date: 04-Dec-1992 #sequence_revision 04-Dec-1992 #text_change 09-Jul-2004
 C:Accession: S22697; S21006
 R:Brtl, H.; Hallmann, A.; Wenzl, S.; Sumper, M.
 EMO J. 11, 2005-2062, 1992
 A:Title: A novel extensin that may organize extracellular matrix biogenesis in Volvox c.
 A:Reference number: S22697; MID:92289669; PMID:1600938
 A:Accession: S22697
 A:Molecule type: mRNA
 A:Residues: 1-464 <HAL>
 A:Cross-references: UNIPROT:Q41645; UNIPARC:UPI000009F7C8; EMBL:X65165; NID:921991; PID
 C:Keywords: glycoprotein

Query Match 11.6%; Score 413; DB 2; Length 464;
 Best Local Similarity 34.5%; Pred. No. 5,5e-12;
 Matches 98; Conservative 27; Mismatches 101; Indels 58; Gaps 12;

QY 30 SASPPAPASANLPKNGQTPPTSDLSBRFASAPP-----PPRRPPPPPTPMPPIAG 85
 DB 229 STSPPPAPVSSSPPPATRSPPRRITSSPSPVLTASPPPKTSPPRRPVSPPPPVAS 288
 QY 86 EPPSPAPASKPTPMPPIAG--BPAPKP-PTPMPPIAGBPAPKPPTPMPPIAGAP 143
 DB 289 PRRPPPVYSPRRPPPPPVYSPRRPPPPPVYSPRRPPPPPVYSPRRPPPPPVYSPRRPPPP 344
 QY 144 TPESQLAPRRPTPTGAPQQPSAPAHVPSHGPHQPRRTAPAPPAKMPGEPAP 203
 DB 345 RPS-----PSRRPSSPRRPPPVYSPRRPPRASBPAPASPPPP--PRPPSPPPPS 397
 QY 204 PSRPSAPAPPPPPAPQHSRRARBGHRYRTDTERVGVKATGSPISQALRAEASGAQL 263

```

Db      398 PPPPTAANPPS-PAPSRSR-----AGCPPL-----423
Qy      264 APGTEPSDAPLGQPRSYLAPTR--PAPTEPPSPSPQNSGR 305
Db      424 --GTRPPPP--PEDDAPPPDYTPPPPDMSPPPKKATGR 462

```

RESULT 9
D96711
hypothetical protein P24J5_8 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
Accession: D96711

C/Accession: D96711
R/Theologias, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chn, C.W.; Hughes, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.P.; Hing, Hughes, B.; Hutzar, L.
Nature 408, 816-820, 2000
A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maltl, R.; Marzall,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shim, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A/Title: Sequence and analysis of chromosome 1 of the Plant Arabidopsis.
A/Reference number: A86141; MUID:21016719; PMID:11330712
A/Accession: D96711
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-708 <5'N>
A/Cross-references: UNIPROT:Q9SKX31, UNIPARC:UPI0000048426, GB:AB005173, NID:95734709, P
C/Genetics:
A/Gene: F24U5.8
A/Map position: 1

Query March	11.3%	Score 402.5;	DB 2,	Length 708;
Best Local Similarity	24.8%	Pred. No. 2.5e-11;		
Matches 179;	Conservative 80;	Mismatches 261;	Indels 203;	Gaps 32;

QY	21	MAAQPFDPDASAPPA-----PASANLPKRNQOTPEPTDLSERVSAPPP-----PPPPP	73
Db	1	MATTPVQPPVNSNPVTSPPPLNANSP--ATPEPTVTSPLP--SAPPEKRAPPPPP	55
QY	74	-----PPPTWMTIAGBP-----PSBPASAKPTPTPMDIAGBPAPKPP	115
Db	56	WTSPPRPVANGAPPPPLPKPPSSSSPPPPQVLPSPPTSSPPPPQVIRSPSPASPPAL	115
QY	116	TPPAPIAGBPAA--PKKPTPMBIAGAPPTPSQLAIPPPTQGTGAPQCPESAP	173
Db	116	VPPSPSSPPPPASVPPPPSPSPPLVNSPPSPSVKPIQSPPPPSDRQTQSPSPSPSP	175
QY	174	HVSHGHPQPKRTAPAPFWAMPIDGEPAP-----SNPSAPAPPTPRAPPOHSRRAR	228
Db	176	--PS--ERFTQSPSPSPSRKPTQSPPPSPSPPSDRSQSPPPPEPDTKQPPRR---	227
QY	229	GHRVTRTERNVGKVAIGPSIQALPKABKAGLAPETES--PAPIGQRYSLAP---	283
Db	228	-----SPNSPPPTSSPPSPREILVPSQN	252
QY	284	-PTRPAPTEPPP--SPSPQRNSG-----RRAERVRHP	312
Db	253	NPQGNNTLPRLDAPNSTNNSSGIGTGAVVISAVVALVFTLPGLPWCIRKRRKRLSA	312
QY	313	DLAAGHAAAPDSTTAATTGGRRRKAAPDIDATOKS-----	349
Db	313	-VSGGDVTPSPMSSTARSDSAPFWQSSAPGASKRSQSYQSGGLGNSKALPSYBELV	371
QY	350	-----LRPAAGPKVKYKVPQ--KPRATPPEKVVSGRWPHW-----VHATLRIR--	392
Db	372	KATNGPQENILGSGFCQCYKGLLPDQRVVAIVAQMLGGGQGDREPAEAVTSLRIHR	431
QY	393	-----LG--LSPEKTELDLHARVRNRGRSGYQIAVGLKGGAGCTTTLTAAGSTLAQVR	445
Db	432	HLVSIYGHGICSGDR--LLIYDVYNNDLVYHNLGKSVLDMATRVKLIAGAAGGLATIH	489

```

QY 446 AD---R1ALADPAGLADRVGRQSGATTADV-LAEKELSHYNDIRAHYSVNAANLEAV 501
Db 490 EDCHEPRIIHRDIKXSNIIIBDNF---DARVSDFGIARLAL---DCNTHITTVIGTGG 541
QY 502 LPABEYSSAOPALSDADWHFIADPASFYNUVLADCGAG---FPDPLTRGVLTSGVSGV 556
Db 542 YMAEPYASGGLTEKSDVF-----SGGVLLBELIGRKPEVDSQPL--GDESLVEMA 591
QY 557 VVAVASVSDGAGQASVALDMLRNNGYOD-----LASRACVYINHMPESEPNVAVVDLVR 610
Db 592 RPLISHAIEETEFPDLADPKLGNYESEMFRMIRBAGCV--RHILATRRPMRG--QIVR 647
QY 611 HFE 613
Db 648 AFE 650

```

RESULT 10
G87008
conserved hypothetical protein M0798 [imported] - Mycobacterium leprae
C1:Species: Mycobacterium leprae
C1:Date: 20-Apr-2001 #sequence revision 20-Apr-2001 #text change 09-Jul-2004

C:Accession: G87008
R: Cole, S.T.; Elgimeier, K.; Parkhill, J.; James, K.D.; Thomson, N.R.; Wheeler, P.R.; H
R., Davies, R.M.; Devlin, K.; Duthoy, S.; Feltsell, T.; Fraser, A.; Hamlin, N.; Holroyd
eam, M.A.; Rutherford, K.M.
Nature 409, 1007-1011, 2001
A: Authors: Rutter, S.; Seeger, K.; Simon, S.; Simmonds, M.; Skelton, J.; Squares, R.; S
A: Title: Massive gene decay in the leprosy bacillus.
A: Reference number: A86909; MUID: 21128732; PMID: 11234002
A: Accession: G87008
A: Status: Preliminary
A: Molecule type: DNA
A: Residues: 1-592 <STO>
A: Cross-references: UNIRROT:Q9CC11; UNIPARC:UPI00000C6D14; GB:AL450280; NID: g13092896;
C: Genetics:
; Gene: ML0798

Query Match	11.0%	Score 391.5	DB 2	Length 592
Best Local Similarity	28.0%	Pred. NO. 6.6e-11		
Matches 129; Conservative	71	Mismatches 201	Indels 59	Gaps 129

```

QY 232 KRTOBERWGVATGSPSIQAALRAEBGAGOLAPETBSPAPLGGPSTYAPPTBPAPPS 291
Db 147 HQTSMKSVLKAAYEPPONSNTDDBNELTDAVALLASTPASHUGOFQALBETTERVSTTN 206
QY 292 PPPSPQPNRSGRAERVRHPDLAAOHA-----AAQPSITATATGRRRKAAP--DUD 344
Db 207 P-----QVADVALLNRWCLRELDDBAAARAL-----AAVPNDP 243
QY 345 ATOKSLPPAAKGPVKVKVKKPKRATKP-----KVSQR-----GRRWVHALTRIN 392
Db 244 TTNHTABAS-----RPQ--PTTPRRPYDDGRDLVARRRRPAGDGRKKVTKATLGR 294
QY 393 LGLSPDEKYEL--DLHARVRBNPGSIOIAYVGLKGAGKTTTLTALCSTLAAQVADRIL 450
Db 295 VNPEPSAKAEBTNEIHRRICAPLADVHKAFVSAKGVGKTTITVALGNTMARLEGRDVI 354
QY 451 ALDDPBAQNLADVRGSG--ATLADVLAKELSHNYDIAHTSVMAVNLVLELRAPBS 508
Db 355 AYVDADADLGDLSAFRRRGGPOTNIHFVSARMAKRYADVVRHTYTMNDRLEMLGQNDP 414
QY 509 SAQALSLDADWHFIADPASFYNLVLADCGAGFPDLFRGLVTSGVVVVSVSIDGAO 568
Db 415 RSTYRLGPDDIETMKLLENHCNVLIDCGIFPVNGPFLNKJVSJTYGLVVAASDDVRGVS 474
QY 569 QASVALDMLBRNGYQDILASRACVYINHIIMPESPUNAVADLVHRHPEQOVOPGRVVMVPMR 628
Db 475 GALATIDMLAHGARLILQHTVVVLNMIQTKKPFVDPBGIVENQPFKRVPD--FYEMPYDP 532
QY 629 HIAAGTEISLDLDPYKRYVLELAALSDFE--RAGR 666

```

Db 533 H1ATG1A5EYTS1KRTKRLKELVGVGAQYTPVIRAGQR 572

RESULT 11

E96636

hypothetical protein T7P1.21 [imported] - Arabidopsis thaliana

C/Species: Arabidopsis thaliana (mouse-ear cress)

C/Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004

C/Accession: E96636

R/Theologus, A.; Ecker, J.R.; Palm, C.J.; Federle, N.A.; Kaul, S.; White, O.; Alonso,

Chin, C.W.; Chung, M.K.; Com, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;

anen, N.F.; Hughes, B.; Hutzler, L.

Nature 408, 816-820, 2000

A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.

C.A.; Li, J.H.; Li, Y.; Liu, S.X.; Liu, Z.A.; Luo, J.S.; Maitl, R.; Marziani,

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shin, P.; Southwick, A.M.; Sun, H.; Tallon,

ker, M.; Wu, D.; Yu, G.; Fraser, C.W.; Venter, J.C.; Davis, R.W.; Sun, H.; Tallon,

A/Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A/Reference number: A86141; MUID:21016719; PMID:11130712

A/Accession: E96636

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1907 <CDS>

A/Cross-references: UNIPROT:Q9C946; UNIPARC:UP100000A9620; GB:AEO5173; NID:96751696; PI

C/Genetics:

A/Map position: 1

Query Match

Best Local Similarity

Matches 137; Conservative 36; Mismatches 127; Indels 208; Gaps 23;

Score 369; DB 2; Length 907;

22 AAGPFPDPSAPPPAPAS-----ANLPKNGCTPPTSDLSERFVSAPPPPPPPPP 77

411 SGLPFP 464

78 TNPPIAAGBPSPBPAPSKPPTPMPPIAGBPBPAPKPPPP--MPPIAGBPAPKPPPP 135

465 AVNPPLKHPAPPP-----PPLPRAVPLKHPAP--PPTPAPPLKHPAPPPPPPP 517

136 MPPIAG--PAPFPSPAPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPP 168

518 TTTAAPPFP 575

189 APPWAKMPFGE-----PPAPSPAPSPAPPPPPPPPPPPPPPPPPPPPPPPPP 242

576 SP--PMPKNSGSGPFPFPFPFPFPFPFPFPFPFPFPFPFPFPFPFPFPFPFP 613

243 VATGPIQARLRAEASGAQLAPCTEFPAPLPQPSYLPAPPPPPPPPPPPPPPPPP 302

614 MANG-----AAGPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPP 642

303 GRAPRRVHPDLAQAAPDSTTAATGGRKRAAPDLDTOK--SLRPAKGPVK 360

643 ---AASLRP-----KKAATLKSTOLGNTYRIKQ--- 671

361 KVPKPKKTKPKPKVVSQGRMHWALTRINGLSPDEKELDLARVVRNPRGSYQA 420

672 KVEGRPNA-----KTSGS----- 686

421 VVGLKGAAGKTTTLAALGSTIAQVRADRLIALDADPGMADRVRGSGATTADYLA 479

687 --GRKAGAGS-----AAGGKQ--MADLALAI 710

480 --KELSHYNDIRAH-----TSVNAVNTLE 501

711 TKKSAVFLQIADIAKMTSINIKIEI 738

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

C/Species: Mycobacterium tuberculosis

C/Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004

C/Accession: A70598

R/Authors: S.T.; Broesch, R.; Parthill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon,

R/Comor, R.; Davies, R.; Devlin, K.; Feltham, T.; Gentles, S.; Hamlin, N.; Holroyd, S.

Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.

Nature 393, 537-544, 1998

A/Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrett, B.G.

A/Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome

A/Reference number: A70500; MUID:98295987; PMID:9634230

A/Accession: A70598

A/Status: preliminary; nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA

A/Residues: 1-341 <COL>

A/Cross-references: UNIPROT:Q05456; UNIPARC:UP100000D0F8A; GB:294121; GB:AL123456; NID:

A/Experimental source: strain H37Rv

C/Genetics:

A/Genes: RV3886C

Query Match

Best Local Similarity

Matches 93; Conservative 66; Mismatches 126; Indels 23; Gaps 7;

Score 386.5; DB 2; Length 341;

372 PPVVSQKGRKRVNALT--RITLGSPPKTELDHARVRNPPSSYQIAYVGLKGA 429

43 PRKIPGTWRKRVVSVSPFKINPGSSPFRHRNLOGRIHRIQYITVSGKGVG 102

430 KTLTAAIGSTIAQVADRILALDADPGAGNLADRVGSGATTADYLAKELSHYNDIR 489

103 VTTMAACIGVRECPENVIALDAVPSFGTLADRIDESPFGYALINDTVGQADIR 162

490 AHTSVANLEVLPADEYSQAQALSDADWHFIADPASPFPYVLVLDCAAGFPDPLTRGV 549

163 EHLGQTVGLDVLVAGNRTSDQFPPLVPMFSAVLSIRBTHTVYIDTSPDLSEHDVMAV 222

550 LSTVSVVYVAVSVISDAQAQASVALDMLRNNGYODLASCAVYINIMGEPRVAYGV 609

223 LOSTDTLVVSGTTPARSRPVLAVDYLAQGHLSVRSYVLNH---TDSITDKDL 278

610 RHP--EQQVPGRVV--VMPMDRHTAAGTEISLDLDLPIY-----KRVLELAALSD--- 658

279 AYITERTVGAIVAEVMPDPHLAKG-----GIIDVHLELNKSRRLRLBITAGLADKY 333

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

QY 271 PAPLQGRSYLAETPRAPTEPPSPSPQNSGRARRVHPDLAAQHAAPDSITAA 330
 DB 14 PHQGTGRGY---SPRPSGDRCPAPPPGR-----HAA-----AAT 46
 QY 331 TGRRRRRRAAPDLDATQSLRPAKGPVKVKKVQKATKP-----PKV-----S 377
 DB 47 PPPPLPSPAP-----LRPLP-----DPMPROQEPAPPPSTMAPALAPISRTSPG 93
 QY 378 ORGMRHVHVALT--RINIGLSPDEKYLDAHVRVRPRGSYOIATVGLKGAGKTTTLTA 435
 DB 94 BRGRRRRVLYTGTGLVGLRSGMORQAQREATITRTYLAHGNKAVLVG-KGGVKTIVAA 152
 QY 436 ALGSTIAQVR-ADRIIALADPAGANLADRVGROSGATTIADVLAKEKLSHYNDIRATSV 494
 DB 153 CVGSIIAHLAQORRIVGIDADTAAGRSLSRIDPPAAGSPWELTTDTMLRSPTDITATLGR 212
 QY 495 NAVNLEPLAPERTSSAORALSDADMHFIADPASFVYLVADCGAGFFDPLTRGVSTVS 554
 DB 213 NSAGLYVLACQAPASGPRRVDLPALYREALRLDHFPAISVYDCSSMEAAVTOELADV 272
 QY 555 GVVVVASVSDGAQASVALDMLNNGYODLASRACVINHMGEBNVAVKDLY-RHPE 613
 DB 273 ALIVVSSPMADGASAAANTTEMUSDYGLTGLLRISIVLVND-SDGHADKRTKSLAQEFT 331
 QY 614 QOYQPGRVVMPMDRHIAAGTEISLD-LDPIYKRYLLELAALSDPF 660
 DB 332 DHQOP--VVEVPDPHLRPGVIMDSHMAFTTRKLQVAAATVATP 377

RESULT 14

P86387
 Probable Pto kinase interactor (imported) - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
 C/Accession: F86387
 R/Theologian, A.; Becker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
 aneas, N.F.; Hughes, B.; Hultzer, L.
 Nature 408, 816-820, 2000
 A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luo, J.S.; Maiti, R.; Marziani,
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shim, P.; Southwick, A.M.; Sun, H.; Tallon,
 Ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A>Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A/Reference number: A86141; MUID:21016719; PMID:11130712
 A/Accession: F86387
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-760 <870>
 A/Cross-references: UNIPROT:Q9C660; UNIPARC:UPI00000A1DD0; GB:AE005172; NID:q11079512; F
 C/Genetics:
 A/Map position: 1

Query Match 10.7%; Score 381.5; DB 2; Length 760;
 Best Local Similarity 23.9%; Pred. No. 2.4e-10;
 Matches 155; Conservative 73; Mismatches 224; Indels 187; Gaps 27;

QY 11 PHEGMEAPDDMAAPPF--DPSAPPPAPASANLPRNGQTPPT-----SDDLSEPF 61
 DB 8 PREBVSISPLASPLMALPPPPQSPFGDNATSPTRTNGNPPETNTNPAQSSPPETP 67
 QY 62 VSAAPPPPPPPPP-----PPTPMPIA-AGRPSPBPASAKPPTPPMIIAGBP---APK 113
 DB 68 LSSPPPPSPSPSLGCGPPPTTIVSPPSPSPPPPTTAPAPPPANVSSPPSPSPPP 127
 QY 114 PPT---PPMIIAGBPAPKPPPTPPMIIAGBPAPPTTSOLAPPR--PPTQTPGAPQ 168
 DB 128 PPTBAPPTTITTSBPPTNPPPPSPSPSLAPDPSPNPLPPCLVPPSPSPHLSPP 187
 QY 169 BSPAPHVSHGPQOPRTAPAPWAKMPIGEPAPAPSPASAPAPPT---PAPQHSR 225

DB 188 ASBIPPPPH-----LPPASERP-STPPSDSHPPPPPPGPKRRBQPPPSGR 238
 QY 226 ARGRHRYTDERNVGVATGSPISQALRAEASGAOLATGBSPAPLQGPSYLAPT 285
 DB 239 PTPSPSPSDSKPVPNPSPPSP-----BETLPPPKVSPPL--PSNSSSPPT 284
 QY 286 --RPAPTEPPSPSPQNSGRARRVHPDLAAQHAAPDSITAAATG----- 332
 DB 285 LPPSSVSPSPSPSPKSVSPDNPSPNNPPTVDNSSSGISGI-AAVAVSIGVALVLT 343
 QY 333 -----GRBRK-----AAPDLATQSLRPAKGPVKVKKVQK 367
 DB 344 LIGVVVCLKKRRKRLSTIGGYVMPPTMESSPSDSA--LLTQSSAPLVGNSSNRT 401
 QY 368 KATPKPK-----VSGRGMRH--VHALTRINIGLSPDEKYLDAH 408
 DB 402 YLQSERGGRGQRELEYSBELVIANTFDENILGEGRGVYKGLPBERV----- 454
 QY 409 VRNPRGSYOIAVGLKGAGK-----TTTLAAGSTIAQVRADRI 451
 DB 455 -----VAVYQLKIGGGQGDREKAEVDITSRVHNNLSMVGYSISERN--RLLI 502
 QY 452 LADPAGANL-----ADRVGROSGAT-----IADVLAE 479
 DB 503 YDYVFN-NMLYFHLAGTPGLDMATRVKIAAGARGLAYHEDCHPRLIHDIKSSNILL 561
 QY 480 KE-----LSHYN-----DIRAHSVNAVNLVLPAPSPSAORALSDAD 518
 DB 562 NNFHALVSDPGLAKLALDNCNTHITTRVMGTGPMABPAYSAGKLTESD 610

RESULT 15

F75518
 Hypothetical protein - Deinococcus radiodurans (strain R1)
 C:Species: Deinococcus radiodurans
 C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
 C/Accession: F75518
 R/White, O.; Bisen, J.A.; Heidelberg, J.F.; Hickey, B.K.; Peterson, J.D.; Dodson, R.J.;
 S.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; M.
 Science 286, 1571-1577, 1999
 A>Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
 A/Reference number: A75250; MUID:20036896; PMID:10567266
 A/Accession: F75518
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-839 <871>
 A/Cross-references: UNIPROT:Q9RX57; UNIPARC:UPI00000C1761; GB:AE001904; GB:AE00513; NI
 A/Experimental source: strain R1
 C/Genetics:
 A/Map position: 1
 C/Superfamily: herpesvirus immediate-early protein IE110; RING finger homology

Query Match 10.6%; Score 377; DB 2; Length 839;
 Best Local Similarity 27.7%; Pred. No. 4.2e-10;
 Matches 152; Conservative 33; Mismatches 221; Indels 142; Gaps 25;

QY 16 EAPDDMAAPF-FDPSAPPPAPASANLPRNGQTPPTSDLSERFVSAPPPPPPPPP 74
 DB 68 EVNNSNAALPLSTSPPLAASPNKAPATKPTPAKP-----ATAPAPPPKPP 116
 QY 75 PPTPMPIAAGRPSPBPASAKP-PPPMPIAGBPAPKPPPTPPMIIAGBPAPPP-P 132
 DB 117 TPPEPKETTPBPCKAPBPCKDPPTABDLKPPVQDTP--PVTPEKVT--PSEVTAPGAP 174
 QY 133 TTP-----MPIA-GPAPPTTSOLAPPPPTP----- 158
 DB 175 TPPEVLAQPPVAQTPVAKPPVAPTSOTPTTPVQAPATRTPPQOARPTENAPAQTPAPA 234
 QY 159 -----QPTPG---APQPSAPPHV--PSHGHPQAPRTAPAPWAKMPIGEP 202
 DB 235 TQAPAAQTPPAQAPATQTPATPAPAAQRPAGAPSP---APAPQANAPAGSVVPEATVP 291

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: April 14, 2006, 17:04:10 ; Search time 245.722 Seconds
(without alignments)
1912.171 Million cell updates/sec

Title: US-10-620-246-70

Perfect score: 3552
Sequence: 1 MAADYDGLFRPHGMEAPDD.....KYLLELAALSDPFERAGRR 666

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : UniProt_05.80.*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3552	100.0	666	Q7TVG2_MYCBO	Q7TVG2 mycobacteri
2	3552	100.0	666	Q69740_MYCTU	Q69740 mycobacteri
3	1550	43.6	586	Q9CDB8_MYCLE	Q9CDB8 mycobacteri
4	1378.5	38.8	478	Q33082_MYCLE	Q33082 mycobacteri
5	752.5	21.2	1083	Q86637_STRCO	Q86637 streptomyc
6	740.5	20.8	810	Q82K60_STRAW	Q82K60 streptomyc
7	680	19.1	525	Q52342_NOCFA	Q52342 nocardia fa
8	655	18.4	446	Q52115_NOCFA	Q52115 nocardia fa
9	638	18.0	594	Q6XN74_RHOBR	Q6XN74 rhodococcus
10	567	16.0	434	Q73SP7_MYCPA	Q73SP7 mycobacteri
11	532.5	15.0	388	Q73VH1_MYCPA	Q73VH1 mycobacteri
12	530	14.9	405	Q8VKJ9_MYCTU	Q8VKJ9 mycobacteri
13	530	14.9	405	Q06396_MYCTU	Q06396 mycobacteri
14	530	14.9	405	Q7U1V4_MYCBO	Q7U1V4 mycobacteri
15	495	13.9	3889	Q6SSR8_CHLNR	Q6SSR8 chlamydomon
16	486	13.7	2371	Q58N45_CHLNR	Q58N45 chlamydomon
17	485	13.7	511	Q95UD0_PIG	Q95UD0 sus scrofa
18	485	13.7	566	Q95JDI_PIG	Q95JDI sus scrofa
19	482	13.6	3409	Q6SSR6_CHLNR	Q6SSR6 chlamydomon
20	478	13.5	409	Q9SBM1_VOLCA	Q9SBM1 volvox cart
21	470.5	13.2	555	GPI_CHLNR	GPI chlamydomon
22	470	13.2	4027	Q512R0_CHLNR	Q512R0 chlamydomon
23	464	13.0	676	Q95UC9_PIG	Q95UC9 sus scrofa
24	462.5	13.0	420	Q9VZC2_DROME	Q9VZC2 drosophila
25	460.5	13.0	436	Q6A844_LEIXX	Q6A844 leifsonia x
26	457.5	12.9	453	Q4NDP0_9MICC	Q4NDP0 bradyrhizob
27	457.5	12.9	745	Q89XK6_BRAJA	Q89XK6 bradyrhizob
28	455	12.8	498	Q5KJ35_CRYNE	Q5KJ35 crypococcu
29	454.5	12.8	1009	Q8L685_VOLCA	Q8L685 volvox cart
30	431	12.1	1627	Q84ZL0_ORYSA	Q84ZL0 oryza sativ
31	430.5	12.1	384	Q745H1_MYCPA	Q745H1 mycobacteri

32	421	11.9	1188	Q41805_MAIZE	Q41805 zea mays (m
33	416	11.7	598	Q8VKN7_MYCTU	Q8VKN7 mycobacteri
34	414.5	11.7	1201	Q9C681_ARATH	Q9C681 arabidopsis
35	414	11.7	760	Q9TOK5_ARATH	Q9TOK5 arabidopsis
36	413	11.6	464	Q41645_VOLCA	Q41645 volvox cart
37	412	11.6	1269	Q7XH56_ORYSA	Q7XH56 oryza sativ
38	412	11.6	1269	Q8W5K6_ORYSA	Q8W5K6 oryza sativ
39	411.5	11.6	1289	Q9FLQ7_ARATH	Q9FLQ7 arabidopsis
40	410.5	11.6	763	Q9XDH2_MYCTU	Q9XDH2 mycobacteri
41	409.5	11.5	3247	Q65553_9ALPH	Q65553 bovine hepr
42	409.5	11.5	3247	Q77CD4_9ALPH	Q77CD4 bovine hepr
43	408.5	11.5	1236	Q9C012_HUMAN	Q9C012 homo sapien
44	408.5	11.5	1250	Q70E73_HUMAN	Q70E73 homo sapien
45	406.5	11.4	687	Q948Y7_VOLCA	Q948Y7 volvox cart

ALIGNMENTS

RESULT 1	Q7TVG2_MYCBO	PRELIMINARY;	PRT;	666 AA.
ID	Q7TVG2_MYCBO	PRELIMINARY;	PRT;	666 AA.
AC	Q7TVG2_			
DT	01-OCT-2003 (TREMBLrel. 25, Created)			
DT	01-OCT-2003 (TREMBLrel. 25, Last sequence update)			
DT	01-OCT-2003 (TREMBLrel. 25, Last annotation update)			
DE	CONSERVED HYPOTHETICAL PROLINE AND ALANINE RICH PROTEIN.			
GN	OrderedAccessionNames=MB3906;			
OS	Mycobacterium bovis.			
OC	Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;			
OC	Corynebacterineae; Mycobacteriaceae; Mycobacterium;			
OC	Mycobacterium tuberculosis complex.			
OX	NCBI_TaxID=1765;			
RN	[1]			
RP	NUCLEOTIDE SEQUENCE.			
RC	STRAIN=AF2122/97;			
RX	MEDLINE=22709107; PubMed=12788972; DOI=10.1073/pnas.1130426100;			
RA	Garnier T., Bigliemeter K., Camus J.-C., Medina N., Mansoor H.,			
RA	Pryor M., Dutkov S., Grondin S., Lacroix C., Monsenpe C., Simon S.,			
RA	Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,			
RA	Parkhill J., Barrell B.G., Cole S.T., Gordon S.V., Hewison R.G.;			
RT	"The complete genome sequence of Mycobacterium bovis."			
RL	Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882 (2003).			
DR	EMBL; BX248347; CAD96092.1; ; Genomic DNA.			
KW	Complete proteome; Hypothetical protein.			
SQ	SEQUENCE 666 AA; 70644 MW; -752E072EB8B10111 CRC64;			
Query Match	100.0%; Score 3552; DB 2; Length 666;			
Best Local Similarity	100.0%; Pred. No. 4.5e-138;			
Matches	666; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	1 MAADYDGLFRPHGMEAPDDMAAOPFPDPASPPAPASANLKPKNQTPPPSDLSER 60			
DB	1 MAADYDGLFRPHGMEAPDDMAAOPFPDPASPPAPASANLKPKNQTPPPSDLSER 60			
QY	61 FVSAPPPPPPPPPPPPPPPPIAAGBPPSEPPASKPTPPMPLAGBPPAPKPTPPMP 120			
DB	61 FVSAPPPPPPPPPPPPPPPPIAAGBPPSEPPASKPTPPMPLAGBPPAPKPTPPMP 120			
QY	121 IAGBPPAPKPTPPMPLAGBPPAPKPTPPMPLAGBPPAPKPTPPMPLAGBPPAPKPTPPMP 180			
DB	121 IAGBPPAPKPTPPMPLAGBPPAPKPTPPMPLAGBPPAPKPTPPMPLAGBPPAPKPTPPMP 180			
QY	181 HQRRRTAPAPPAKMPIGBPPAPSPRSASAPAPPPPPAPAPPPPPAPAPPPPPAPAPPPPP 240			
DB	181 HQRRRTAPAPPAKMPIGBPPAPSPRSASAPAPPPPPAPAPPPPPAPAPPPPPAPAPPPPP 240			
QY	241 GKATGTSIARARASASGAOLAPGTPSPAPLGGORSYLAAPTPAPTPAPTPAPTPAPTPAPTP 300			
DB	241 GKATGTSIARARASASGAOLAPGTPSPAPLGGORSYLAAPTPAPTPAPTPAPTPAPTPAPTP 300			
QY	301 NSGRRAERRVHPDLAAOHAAPDSITTAATYGGRRRRAAPDDATOKSLRPAKGBKVK 360			
DB	301 NSGRRAERRVHPDLAAOHAAPDSITTAATYGGRRRRAAPDDATOKSLRPAKGBKVK 360			

```

Db      301 NSGRRAERRVHDLAAQHAAPDSITTAATTGRRRRRAAPDLDTOKSLRPAKGPVK 360
QY      361 KYVPOKPKATKPKPVVSQSGMRHWALTRINLGLSPDEKYHDLHARVRNPRGSYQIA 420
Db      361 KYVPOKPKATKPKPVVSQSGMRHWALTRINLGLSPDEKYHDLHARVRNPRGSYQIA 420
QY      421 VVGLKGAAGKTTLTAAAGSTLAQVRADRIALDADPGAGNLADVRGSGATTADYLAEX 480
Db      421 VVGLKGAAGKTTLTAAAGSTLAQVRADRIALDADPGAGNLADVRGSGATTADYLAEX 480
QY      481 ELSHYNDIARHTSVNAVNLVLPAPRYSSAQRALSDADWHFIADPASRFYNVLADCGAG 540
Db      481 ELSHYNDIARHTSVNAVNLVLPAPRYSSAQRALSDADWHFIADPASRFYNVLADCGAG 540
QY      541 FPDPLTRGVLTSTVSGVVAASVSDGAQOASVALDMLRNNGYODLASRACVYNHIMPGE 600
Db      541 FPDPLTRGVLTSTVSGVVAASVSDGAQOASVALDMLRNNGYODLASRACVYNHIMPGE 600
QY      601 PNVAVDLVRHFEQVQVQPGRVVMPMDRIIAAGTEISLDLDPYKRKVLAAALSDF 660
Db      601 PNVAVDLVRHFEQVQVQPGRVVMPMDRIIAAGTEISLDLDPYKRKVLAAALSDF 660
QY      661 ERAGRR 666
Db      661 ERAGRR 666

RESULT 2
ID      069740 MYCTU PRELIMINARY; PRT; 666 AA.
AC      069740; Q7D4P3;
DT      01-AUG-1998 (TrEMBLrel. 07, Created)
DT      01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT      13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE      CONSERVED HYPOTHETICAL PROLINE AND ALANINE RICH PROTEIN (Hypothetical
        protein).
GN      OrderedLocusNames=MT3990, RV3876;
OS      Mycobacterium tuberculosis.
OC      Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC      Corynebacteriaceae; Mycobacteriaceae; Mycobacterium;
OC      Mycobacterium tuberculosis complex.
OX      NCBI_TaxId=1773;
RN      [1]
RP      NUCLEOTIDE SEQUENCE.
RC      STRAIN=H37R;
RX      MEDLINE=98295987; PubMed=9634230; DOI=10.1038/31159;
RA      Cole S.T., Broch R., Parkhill J., Garnier T., Churcher C.M.,
RA      Harris D.E., Gordon S.V., Eigmeier K., Gas S., Barry C.E. III,
RA      Tekala F., Badcock K., Basham D., Brown D., Chillingworth T.,
RA      Connor R., Davies R.M., Devlin K., Feltham T., Gentles S., Hamlin N.,
RA      Holtroyd S., Hornsby T., Jagels K., Krogh A., McLean J., Moule S.,
RA      Murphy L.D., Oliver S., Osborne J., Quail M.A., Rajandream M.A.,
RA      Rogers J., Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA      Sulterson J.E., Taylor K., Whitehead S., Barrett B.G.;
RT      "Deciphering the biology of Mycobacterium tuberculosis from the
RT      complete genome sequence.";
RL      Nature 393:537-544(1998).
RN      [2]
RP      NUCLEOTIDE SEQUENCE.
RC      STRAIN=CDC 1551 / Oshkosh;
RX      MEDLINE=22206494; PubMed=12218036;
RX      DOI=10.1126/JB.184.19.5479-5490.2002;
RA      Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA      Peterson J.D., Deboy R.T., Dodson R.J., Gwinn M.L., Haft D.H.,
RA      Hickey E.K., Kolonay J.F., Nelson W.C., Umayam L.A., Esmolaeva M.D.,
RA      Salzberg S.L., Delcher A., Utterback T.R., Welldman J.F., Knout H.M.,
RA      Gill J., Minkula A., Bishai W., Jacobs W.R. Jr., Venter J.C.;
RA      Fraser C.M.;
RT      "Whole-genome comparison of Mycobacterium tuberculosis clinical and
RT      laboratory strains.";
RL      J. Bacteriol. 184:5479-5490(2002).
DR      EMBL; BX842584; CAI17968.1; -; Genomic DNA.
DR      EMBL; AE000516; AAK48358.1; -; Genomic DNA.

```

```

DR      PIR; B70803; B70803.
DR      TIGR; MT3990; -.
DR      Tuberculist; RV3876; -.
KW      Complete proteome; Hypothetical protein.
SQ      SEQUENCE 666 AA; 70644 MW; 752E072F8B8B1011 CRC64;

Query Match      100.0%; Score 3552; DB 2; Length 666;
Best Local Similarity 100.0%; Pred. No. 4.5e-138;
Matches 666; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MAADYKLRPRHGMKAPDDMAAQPDPDSAPSPAPASANLPKPGQTPPPPSDLSR 60
Db      1 MAADYKLRPRHGMKAPDDMAAQPDPDSAPSPAPASANLPKPGQTPPPPSDLSR 60
QY      61 FVSAPPPPPPPPPPPPTMPPIAAGSPSPSPASAPPTPMPPIAGSPBPAPKPTPPMP 120
Db      61 FVSAPPPPPPPPPPPPTMPPIAAGSPSPSPASAPPTPMPPIAGSPBPAPKPTPPMP 120
QY      121 IAGBPAPKPPPTPMPPIAGPAPPTPESQLABPPPTPOTPTGAPQOPSPAPPHVSHGP 180
Db      121 IAGBPAPKPPPTPMPPIAGPAPPTPESQLABPPPTPOTPTGAPQOPSPAPPHVSHGP 180
QY      181 HQPRRTAPAPPAKXMPISGPPAPSPSPASAPSPPTPAPQHSRRARRGHRTOTERNV 240
Db      181 HQPRRTAPAPPAKXMPISGPPAPSPSPASAPSPPTPAPQHSRRARRGHRTOTERNV 240
QY      241 GKVAATGPSIQARLRAEASGAQLAPGTEPSAPLPGPSPSYLAPPTPAPTEBPPSPSPOR 300
Db      241 GKVAATGPSIQARLRAEASGAQLAPGTEPSAPLPGPSPSYLAPPTPAPTEBPPSPSPOR 300
QY      301 NSGRRAERRVHDLAAQHAAPDSITTAATTGRRRRRAAPDLDTOKSLRPAKGPVK 360
Db      301 NSGRRAERRVHDLAAQHAAPDSITTAATTGRRRRRAAPDLDTOKSLRPAKGPVK 360
QY      361 KYVPOKPKATKPKPVVSQSGMRHWALTRINLGLSPDEKYHDLHARVRNPRGSYQIA 420
Db      361 KYVPOKPKATKPKPVVSQSGMRHWALTRINLGLSPDEKYHDLHARVRNPRGSYQIA 420
QY      421 VVGLKGAAGKTTLTAAAGSTLAQVRADRIALDADPGAGNLADVRGSGATTADYLAEX 480
Db      421 VVGLKGAAGKTTLTAAAGSTLAQVRADRIALDADPGAGNLADVRGSGATTADYLAEX 480
QY      481 ELSHYNDIARHTSVNAVNLVLPAPRYSSAQRALSDADWHFIADPASRFYNVLADCGAG 540
Db      481 ELSHYNDIARHTSVNAVNLVLPAPRYSSAQRALSDADWHFIADPASRFYNVLADCGAG 540
QY      541 FPDPLTRGVLTSTVSGVVAASVSDGAQOASVALDMLRNNGYODLASRACVYNHIMPGE 600
Db      541 FPDPLTRGVLTSTVSGVVAASVSDGAQOASVALDMLRNNGYODLASRACVYNHIMPGE 600
QY      601 PNVAVDLVRHFEQVQVQPGRVVMPMDRIIAAGTEISLDLDPYKRKVLAAALSDF 660
Db      601 PNVAVDLVRHFEQVQVQPGRVVMPMDRIIAAGTEISLDLDPYKRKVLAAALSDF 660
QY      661 ERAGRR 666
Db      661 ERAGRR 666

RESULT 3
ID      09CDD8 MYCLE PRELIMINARY; PRT; 586 AA.
AC      09CDD8;
DT      01-JUN-2001 (TrEMBLrel. 17, Created)
DT      01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT      01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE      Hypothetical protein ML0048.
GN      OrderedLocusNames=ML0048;
OS      Mycobacterium leprae.
OC      Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC      Corynebacteriaceae; Mycobacteriaceae; Mycobacterium;
OX      NCBI_TaxId=1769;
RN      [1]

```

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=TN;
 RX MEDLINE=2118732; PubMed=11234002; DOI=10.1038/35059006;
 RA Cole S.T., Eigmeier K., Parkhill J., James K.D., Thomson N.R.,
 RA Wheeler P.R., Honore N., Garnier T., Churcher C.M., Harris D.E.,
 RA Mungall K.L., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Davies R.M., Devlin K., Duthoy S., Feltwell T., Fraser A., Hamlin N.,
 RA Holroyd S., Hornby T., Jagsis K., Lacroix C., Maclean J., Moule S.,
 RA Murphy L.D., Oliver K., Quail M.A., Rajandream M.A., Rutherford K.M.,
 RA Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R.,
 RA Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,
 RA Barrett B.G.;
 RT "Massive gene decay in the leprosy bacillus";
 RL Nature 409:1007-1011 (2001).
 DR EMBL; AL583917; CAC29556.1; -; Genomic_DNA.
 DR PIR; H86914; H86914.
 DR Leptoma; ML0048; -;
 KW Complete proteome; Hypothetical protein.
 SQ SEQUENCE 586 AA; 63226 MW; 0FDALB54F36D69DA CRC64;

Query Match 43.6%; Score 1550; DB 2; Length 586;

Best Local Similarity 49.2%; Pred. No. 4.4e-56;
 Matches 335; Conservative 66; Mismatches 168; Indels 112; Gaps 15;

QY 1 MAADYDGLFRPHGMEAPDDMAQPFDPSPAPASANTLPKXGQTPPTSDLSER 60
 DB 1 MAADYDGLFRPHGMEAPDDMAQPFDPSPAPASANTLPKXGQTPPTSDLSER 59
 QY 61 FVAPAPPP 120
 DB 60 L-----PPNPAPASKPLPMPPI--GSSVQPPPASSESPAPMPVASP---PRSPASLMP 110
 QY 121 IAGPEAPAP-----KPPTPMPIAGAPAPPTPSQOLAPRPPPTPOTPGAPQPR 169
 DB 111 ISEPPQPPPEAPABHQAAPSPSVPIPINEPSP-----AKATPMPTPIDSGSR-- 161
 QY 170 SPAPHVSHGPHQPRRTAPAPPAKMPIGBP--PAPS---RPSASAPPPTRPAPQSR 224
 DB 162 TPV-----TSPPSLAFPS-AQPAPATPKSLPAPMSPPPEAPRPSANQHSR 207
 QY 225 RARRGRYRTDTERNVKATGTSIQARLAEBSGAQLAPGTSPAPLGPQPSYLAAP 284
 DB 208 HARRGHHHDETQANPASATEPMIAPARTAE---LRQAPHAALAPAP----- 253
 QY 285 TRAPTEPPSPSPQNSGRAERVRVPDLAQAHAQPSITPAATGGRRRKRAAPDLD 344
 DB 254 -----TQHLTRPDGLVSHRTALHDS-TATSAIGVQYGRST----- 287
 QY 345 ATQKSLRPAKGPVKVKKQPKPATKPPKVSQGRHWHVHALTRINLGLSPDEKYLE 404
 DB 288 -----GAKKPSKVAAKGRHWHVHTYTRINLGLSPDERYEL 324
 QY 405 LHAARVRNPRGSYQIAYVGLKGAGKTTTLAALGSTLAQVRADRIALDADPGANLADR 464
 DB 325 LRTVRVRNPRGSYQIAYVGLKGAGKTTTLAALGSTLAQVRADRIALDADPGANLADR 384
 QY 465 VGRQSGATTADYLAKEKLSHYNDIRATSVVAANLVTLPAPRYSQAQALSDADMHFIAD 524
 DB 385 AGFSSRANLADLADKVSINDIRATSVVAANLVTLPAPRYSQAQALSDADMHFIAD 444
 QY 525 PASRPNLVLADGAGFPDPLTRGVLTSGVSVVVAASVSDGAQOASVALDMLNNNGYOD 584
 DB 445 TVSKYNNVMLADGAGFPDPLTRGVLTSGVSVVVAASVSDGAQOASVALDMLNNNGYOD 504
 QY 585 LASRACVYVNHIMPGEENVAVKDILVRHEQOVQGRVVVMPMDRIIAAGTISLIDLPI 644
 DB 505 LLSRACVYVNHIMPGEENVAVKDILVRHEQOVQGRVVVMPMDRIIAAGTISLIDLPI 564
 QY 645 YKXKVLTAALASDDPERAGR 665
 DB 565 YKXKVLTAALASDDPERAGR 585

RESULT 4

ID 033082 MYCLE PRELIMINARY; PRT; 478 AA.
 AC 033082;
 DT 01-JAN-1998 (TReMBLrel. 05, Created)
 DT 01-JAN-1998 (TReMBLrel. 05, Last sequence update)
 DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
 DE Hypothetical protein MLCB628.11c.
 GN Name=MLCB628.11c;
 OS Mycobacterium lepreae.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxID=1769;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=93188700; PubMed=8446027;
 RA Eigmeier K., Honore N., Woods S.A., Caudron B., Cole S.T.;
 RT "Use of an ordered cosmid library to deduce the genomic organization
 of Mycobacterium lepreae";
 RL Mol. Microbiol. 7:197-206 (1993).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.

QY 95 SKRPTPMPIAGSEP--APKPPPTPMPIAGSEPAPKPPPTPMPIAGAPPTPSQOLAP 152
 DB 4 SEPPQPPPEAPABHQAAPSPSVPIPINEPSPAPATPMPTPIDSGRTPTVS--- 58
 QY 153 PREPTQTPGAPQSPAPHVSHGPHQPRRTAPAPPAKMPIGBPAPSPASPA 212
 DB 59 ---PEPSLAFPS-AQPAPATPKSLPAPMSPPPEAPRPSANQHSR 95
 QY 213 EPPTRAPQHSRARRGRYRTDTERNVKATGTSIQARLAEBSGAQLAPGTSPAPLGPQPSYLAAP 272
 DB 96 -----QHSRARRGHHHDETQANPASATEPMIAPARTAE---LRQAPHAALAPAP----- 144
 QY 273 ELGQPSYLAAPTRPAPTEPPSPSPQNSGRAERVRVPDLAQAHAQPSITPAATG 332
 DB 145 P-----TQHLTRPDGLVSHRTALHDS-TATSAI 171
 QY 333 GRRKRAAPDLADTQKSLRPAKGPVKVKKQPKPATKPPKVSQGRHWHVHALTRIN 392
 DB 172 GVQYGRST-----GAKKPSKVAAKGRHWHVHTYTRIN 204
 QY 393 LGLSPDEKYLELHAARVRNPRGSYQIAYVGLKGAGKTTTLAALGSTLAQVRADRIAL 452
 DB 205 LGLSPDERYELDRTVRVRNPRGSYQIAYVGLKGAGKTTTLAALGSTLAQVRADRIAL 264
 QY 453 DADPGAGNLADRGSGATYADYLAKEKLSHYNDIRATSVVAANLVTLPAPRYSQAQALSDADMHFIAD 512
 DB 265 DADTSCGNLADRGSGATYADYLAKEKLSHYNDIRATSVVAANLVTLPAPRYSQAQALSDADMHFIAD 324
 QY 513 ALSDADMHFIADRGSGATYADYLAKEKLSHYNDIRATSVVAANLVTLPAPRYSQAQALSDADMHFIAD 572
 DB 325 ALSGEDMNPAAAVSVKTYNNVMLADGAGFPDPLTRGVLTSGVSVVVAASVSDGAQOASV 384
 QY 573 ALDMLNNNGYODLASRACVYVNHIMPGEENVAVKDILVRHEQOVQGRVVVMPMDRIIAA 632
 DB 385 ALDMLNNNGYODLASRACVYVNHIMPGEENVAVKDILVRHEQOVQGRVVVMPMDRIIAA 444
 QY 633 GTTISLIDLPIYKXKVLTAALASDDPERAGR 665
 DB 445 GTTIRLRLDPLYKXKVLTAALASDDPERAGR 477

```
RESULT 5
086637 STRCO PRELIMINARY; PRT; 1083 AA.
ID 086637 STRCO
AC 086637
DT 01-NOV-1998 (T-EMBLrel. 08, Created)
DT 01-NOV-1998 (T-EMBLrel. 08, Last sequence update)
DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)
DE Hypothetical protein SC05717.
GN OrderedLocusNames=SC05717; ORFNames=SC3C3.03c;
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycinae; Streptomycetaceae; Streptomycetes.
OX NCBI_TaxID=1902;
RN (1)
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=A3(2) / M145;
RX MEDLINE=21996410; PubMed=12000953; DOI=10.1038/417141a;
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kleeer H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kleeer T., Larke L., Murphy L.D., Oliver K., O'Neill S.,
RA Rabinovitch E., Rajandream M.A., Rutherford K.M., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J.R., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
coelicolor A3(2)."
RL Nature 417:141-147(2002).
DR EMBL; AL93124; CAZ0252.1; -; Genomic DNA.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 1083 AA; 112083 MW; AD42EDF64C12126D CRC64;
```

```
Query Match 21.2%; Score 752.5; DB 2; Length 1083;
Best Local Similarity 26.5%; Pred. No. 3.7e-23;
Matches 252; Conservative 90; Mismatches 253; Indels 357; Gaps 31;

QY 13 EGMEDPDDMAQPFPPDSASFPAPASANLKP-NGQTPPTSDLSERFVSAP----- 65
DB 56 EGMADPPGGSAG-----SPPASAPPAAPAPAPAPAPAPAPAPAPAPAPAPAPAPAP 111
QY 66 --PPPP-----PPPTP-----MPIAGBPSPSPA-ASKP 97
DB 112 FTTPPPAPAPAGAAFTPPPPPGAPPTPPAPPGSGPAGVXKLPVSGSFEPQSPORSAP 171
QY 98 PTP-----PMIAGBPAP-----PKRPTP-----PMPTA 122
DB 172 PEBEPPEVPTPEPEPKAAVAPAPADLASPAALPSFPAGGQLOAPVPPPPAEEFGAA 231
QY 123 GEPAP----- 128
DB 232 EEPPEDEPRDPTDGLDNGDLSGATMRFSAVAKRIAPAAADPTAGSPPAGADBEAD 291
QY 129 ----- 128
DB 292 TPDEPFTVGTPTDADTDAETDVDTESDPESVSDQAPDAESGGHGVADPRLGSAADAE 351
QY 129 --PKRPTPMPITAGAPPTRES-----QLAP 152
DB 352 PEBEPPESEPVVLAAVAESSESERVGVDEDEDSVPTADVDAVPDDAAPADSVQDAV 411
QY 153 PRPPTQTPTGAPQCESPAHPVS---HGPHQR-----RTAPAP-----WAKMPIGE 199
DB 412 PQDATTAPAEAGMPPLPSPSYFVPAALAAHQMPPAPQPPPTQITAPAPPEQGVAAQPPPG 471
QY 200 ---PPAPASPSAPAP-----TRPAQHSRRARGHRYTDTENVKQAT 245
DB 472 PQAPQAPAPAGMDAPAPAPQPGYFPPQGAQAPQOPHSPAPPGYGF-----PQAPQAPH 526
QY 246 GPGITQ-----AKLAERASGAQLAPGTTP-SPAPLQGP-----RSTLAP 284
DB 527 GSPFQOQPPYGVVQOQAPQAGGGOAPAPAPQPGAPQGPWPQGPQSGQPGQAPSAAPQ 586
```

```
QY 285 TRPAPTEPPS-----PSPQNSGRR-AERV-----HPDLAQHAA 321
DB 587 AABNPEQAPPAQOPPEVQAPQOPQOPQOPQOPQOPQOPQOPQOPQOPQOPQOPQOPQ 646
QY 322 QPDSTIAP-----TGRRRRRRAAPDLDATQKSLRPAKQPKVKKVQPKKATKPPK 374
DB 647 APLGTAAVLEBLSDRLVNNKRRQ-----AKSRPAAAG----- 679
QY 375 VVSQGRMVMVALTRINGLSPDE---KYELDLHARVVRNPGSYQIYAVGLKGAGKT 431
DB 680 -----SRFKLGKKEERQRKDL---ITPVLSCRIAVISLKGVGKT 722
QY 432 TLTALGSTLAQVRADRIALDADPGAGVLADRVGQSGATTAD-VLAKEKSHYNDIYA 490
DB 723 TTTALGSTLATERQKILAIIDANPAGTIGRRVRETAGTTRDLVQAPYLINSYMDIR 782
QY 491 HRSVNAVNLVLPAPRYSQAQALSPADNHFPLADPRSRVYNVLTDCGAFPDPLRGVL 550
DB 783 FTSQASSGLEITANDVDPVSTTFNEDYRRAIDVIGQOYPTILDSGTGLYSANRGVL 842
QY 551 STVSGVVVAVSYIDGAQOASVALDWLRNNGYQDLSRACVYINHMPEEPVNAVVDLYR 610
DB 843 DLADQIIITSTSVDCASSASTTLDWLSAHGYADLVSRGITVYISGVRGKMKVDDIYG 902
QY 611 HFEQOVQGRVVMPPDRHIAAGTEISLDLDPYRKVLBELAALSDPFR 662
DB 903 HFQTRCR--GVVVVPFDEHLSAGAEVLDLDMRPKYAEAVFNLAAMVAEDFVR 952
```

```
RESULT 6
082K60 STRAW
ID 082K60 STRAW PRELIMINARY; PRT; 810 AA.
AC 082K60;
DT 01-JUN-2003 (T-EMBLrel. 24, Created)
DT 01-JUN-2003 (T-EMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=SAV2544;
OS Streptomyces avermitilis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycinae; Streptomycetaceae; Streptomycetes.
OX NCBI_TaxID=33903;
RN (1)
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX MEDLINE=22608306; PubMed=12692562; DOI=10.1038/nbt820;
RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
RA Sakaki Y., Hattori M., Omura S.;
RT "Complete genome sequence and comparative analysis of the industrial
microorganism Streptomyces avermitilis."
RL Nat. Biotechnol. 21:526-531(2003).
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX MEDLINE=21477403; PubMed=11572948; DOI=10.1073/pnas.211433198;
RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osone T.,
RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
RT "Genome sequence of an industrial microorganism Streptomyces
RT avermitilis: deducing the ability of producing secondary
RT metabolites."
RT Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
DR EMBL; BA000030; BAC70255.1; -; Genomic DNA.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 810 AA; 86244 MW; F8D19CDA203FDC26 CRC64;
```

```
Query Match 20.8%; Score 740.5; DB 2; Length 810;
Best Local Similarity 31.0%; Pred. No. 8.8e-23;
Matches 217; Conservative 72; Mismatches 246; Indels 165; Gaps 22;

QY 8 LRRPHEGMAFPDDMAQPFPPDSASFPAPASN-----NLKPRGCTPPPTSDLSRFTS 63
```

Db	11.8	NRPRPRLPC---	RRRP---	TPRARTGAPAPAPQAGGFPQGPVTPPPAPADSNVKDQDYG	1172		
Qy	64	APPPPPP	PPPPPPPTMPPLIAGBPSSBPAAKSPPTPMPPLIAGBPAPKCP	-----	1155		
Db	173	FPQAGAGVFPQAGFPAPGPISPAAQAPAP	PAPAPAPPTPQPGYFPQPGFPVPQSGYGF	2311			
Qy	116	----	TPPMLIAPBPAPKCP	-----	PTPMPPLIAGAPPTTBSQALAPRPPTPTQTPGAP	1655	
Db	232	POQGAAPBPAPAPQDPBPAPQAGYGFPBPAPAPAG	-----	DAHNPPTQPGYGF	2822		
Qy	166	QPPSPAPHPVPHSGHPQAPRTAPAPWAKMFGPBPAPSPAPSPAPBPPTBPAPQHSRR	2255				
Db	283	QCGGFPAPQAGCGFPBPAPQAHQAP	-----	CGP	-----	QGRAVPAPAAAPH	3244
Qy	226	ARRGHRRTDTERNVGKATGSPSIQALRAEASGAQLAPGTPSPAPILQPSYSIAPPT	2855				
Db	325	----	PG	-----	PQSGYGF	PATPPPPA	3434
Qy	286	RPA	-----	PTPPPPSPSPKNSGRARVRVHDLAA	-----	QRAAPDSITATTC	3322
Db	344	PPAQAPAPAPQSPQFPQFPQ	-----	PPVDPRTGAAMPQPVQHDQRPPTNPVAPILG	3959		
Qy	333	GRRRGAAPDLATQ	-----	KSLRPAKAPKPKKVPQPKATKPKPVVVSQRGWRH	3633		
Db	396	----	YTPAVELSSDRILNKKQKAKSGRGGA	GGGLFKLGAKKGBEAR	-----	4388	
Qy	384	VWHALTRINLGLSPDEKVELDLHARVRNPRGSGYQIAYVGLKGGAGKTTTLTAAGSTLQ	4433				
Db	439	----	QRKEL	-----	IRTPVLSCYRLAVISLKGVGKTTTLTALGATLAT	4797	
Qy	444	VRADRIILALDADPGAGNLADRVGRGSGATITAD	VLAKELSHNDIRATISVNAVNLVL	5022			
Db	480	ERODKTLIDANPDAGTLGRVRREKATIRDLVOAIPYLNSTMDIRRPFSQAPSGLEII	5399				
Qy	503	PAPERSAQBALSDADWHFIADPASRFRVYLVDGAGFEDPLTRGVLTVSGVAVVAVS	5622				
Db	540	ANDVDPASTTYNDESDYRAIDVLGKQFPILITLSDSTGLYSAMRGVLDLADQILITIP	5999				
Qy	563	SIDGAQASVALDMLNNQGYODLASRACVIVNINIMGPBPVAVKDLVRRHPQOVQBPGRV	6222				
Db	600	SYDGASSASTTLDWLSAHGYADLVNSITVIVSGVRRTGKTIKYEDIVGHFETCR	6577				
Qy	623	VMPWDRHIIAAGTBSLIDLDPYKKRKVLELAALASDPFER	6622				
Db	658	VVPFDEHIIAAGARVLDLMMRPKVREAYFNILSVVAEDMAR	6977				
RESULT 7							
Q52342	NOCFA	PRELIMINARY;	PRT;	525 AA.			
AC	052342						
DT	25-OCT-2004	(TREMblrel. 28, Created)					
DT	25-OCT-2004	(TREMblrel. 28, Last sequence update)					
DT	25-OCT-2004	(TREMblrel. 28, Last annotation update)					
DE	Hypothetical protein.						
GN	OrderedLocustNames=nfa3070;						
OS	Nocardia farcinica.						
OC	Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;						
OC	Corynebacterineae; Nocardiaceae; Nocardia.						
OX	NCBI_TaxID=37329;						
RA	NCUIEOTIDE SEQUENCE.						
RC	STRAIN=IFM 10152;						
RX	PubMed=15466710; DOI=10.1073/pnas.0406410101;						
RA	Ishikawa J., Yamashita A., Mikami Y., Hoshino Y., Kurita H., Hotta K.,						
RA	Shiba T., Hattori M.,						
RT	"The complete genomic sequence of Nocardia farcinica IFM 10152."						
RL	Proc. Natl. Acad. Sci. U.S.A. 101:14925-14930(2004).						
DR	EMBL; AP006018; BAD55149.1; -; Genomic_DNA.						
DR	InterPro; IPR001945; ATPase_a/centrue_DNA.						
DR	PROSITE; P800152; ATPASE_ALPHA_BETA; UNKNOWN_1.						
DR	Complete proteome; Hypothetical protein.						
QO	SEQUENCE 525 AA; 55760 MW; F404F846B85AA09D CRC64;						

Query Match	19.1%;	Score 680;	DB 2;	Length 525;
Best Local Similarity	28.7%;	Pred. No. 1.8e-20;		
Matches 194;	Conservative 66;	Mismatches 205;	Indels 212;	Gaps 16;
Qy	18	PDDMAAOP-----	FFDPSASF-PAPASAMLPKFN-----	GQTPPTSLSLRF 61
Db	18	PTDTRABEGEASATARTPEAGBPAPBAAAGFAPGCFEGAGRAPDGFAPYSGQF		77
Qy	62	VSADPP-----	PPPPPPPPPTMTPLAAGPSPSPBPAAKSPPTPTMTPLAAGPBPAPRP	115
Db	78	AQAGAPFGGAPRPPPPQGPSPAPSYBYRQZBPDGVNRVPPDGAEGVPOS-----		132
Qy	116	TPPMPIAGPSP-----	APPKPTPTMTPLAAGPAPPTTESQLAPRPPTPTPTGAPQGPSP	171
Db	133	-----TGQGSYMAAPPPPTTAMP-----	PPQPPMTQGPAGAP-----	167
Qy	172	APHVSHGPPHQPRTTAPAPPAKMP--	IGBPAPSPBSAPAPBPPTTAPDQSHRARG	229
Db	168	-----	PGQPPGPPQYGGPQGGPQGGPQGGP--	PQGY----- 201
Qy	230	HRVYTDTRBNVKGATGSPISQARLAEASAGQLAPTEBPAPLGGPRTSLAPPTRAP		289
Db	202	-----	GGPMM-----	206
Qy	290	TEPPSPSPQSNRGRABRRVHPLAQAHAAPDSITATTTGRRRRKAAPDLDAQKS		349
Db	207	--PFGHSVNDLMLKRRP-----	AP-----	225
Qy	350	LRPAKGPQKVKVPQKPKATPKPVVSQGRMNHVALTR--	INLGLSPDEKYLDDHA	407
Db	226	-----	RSGMRAVYKAASGSIINPGESADIVYRDLYE	257
Qy	408	RYRKNRPGSYQIAVVGKLGAGKTTTLPAAGSTLQAQRADYILMLDDPGAGNLADRYGR		467
Db	258	RVNPPGVGDYRIALISLKGVGKTTTGVGASTPASIARGDVIAIDANPDGLTAHRVPR		317
Qy	468	QSGATVADVLAEKLSHYNDIRATSYVNAVVLGYLPAPETSSAQRALSDADMHFIADBPAS		527
Db	318	QTRTATKNLEDDQIISKSYDVRATTSQAPSLATSLASBQDPAVEABPSBADYRKAIGILO		377
Qy	528	RFYNVLVADCGAGFPDPLTRGVLSVSGVVVVASVSDGAQAQASVALDMLRNNGYDIAS		587
Db	378	SFYNIILTDGSGTGLMHSAMAGVLDMASSLVLTSPALDAGASASATLDMLEHHGYSKLYE		437
Qy	588	PACVYINIMGSRNNAVKQLYVRFPEQOVQGRVYVNPMPNHTAAGTEISLDDLPYKR		647
Db	438	RTVVVVNARSRGASTVLDOLRKLFLERTR--	AVQVPPFDHILABGAEIDLELVSKEPTRR	495
Qy	648	KVLEIAAALSDDFERRAG	664	
Db	496	ALBELAAMVADDFGYG	512	
RESULT 8				
OSZLIS.NOCFA PRELIMINARY; PRT; 446 AA.				
AC	OSZLIS_	25-OCT-2004	(TRENBLrel. 28, Created)	
DT	25-OCT-2004	(TRENBLrel. 28, Last sequence update)		
DT	25-OCT-2004	(TRENBLrel. 28, Last annotation update)		
DE	Hypothetical protein.			
GN	OrderedLocustNames=nta8310;			
OS	Nocardia farcinica.			
OC	Bacteriota; Actinobacteriota; Actinobacteridae; Actinomycetales;			
OC	Corynebacterineae; Nocardiaceae; Nocardia.			
OX	NCBI_TaxID=37329;			
RN	[1]			
RP	NUCLEOTIDE SEQUENCE.			
RC	STRAIN=IRM 10152;			
EX	PubMed=15466710; DOI=10.1073/pnas.0406440101;			
RA	Tshikwa U., Yamashita A., Mikami Y., Hoshino Y., Kurita H., Hotta K., Shiba T., Hattori M.;			

DR EMBL; AB017241; AAS06576.1; -; Genomic_DNA.
 KW Complete proteome; Hypothetical protein.
 SQ SEQUENCE 434 AA; 46331 MM; 73A47808AD6D13B1 CRC64;

Query Match 16.0%; Score 567; DB 2; Length 434;
 Best Local Similarity 31.3%; Pred. No. 6.3e-16;

Matches 162; Conservative 68; Mismatches 185; Indels 102; Gaps 15;

QY 158 PGTPTGAPQOPEBPAPVSHGHQPRRTAPAPWAKPITGPPAPASPSASAPPT 217
 DB 5 PTGAGVAPBEQOTQIIPAAAAGDE-----KQDAAPQDPBP-----GDAPT- 45
 QY 218 PAPDSRRARRGRHRTDTERNVKVAATGPSIOARLRAEASGAOLAPGTEPSAPLQOP 277
 DB 46 -----KAPAGFR-----TERRV-----PAPREP 64
 QY 278 RSVIAPRTAPRTPEPPSPS-----PQNSGRABRRVHEDLAAQHAAPDSITATT 331
 DB 65 ----APPTAPPPGMPMDSTPTVTGTP-----RVDEPTAGATYAG--PDAPPTQIQ 109
 QY 332 GGRRRKKAAPDLDTQKSLRPAAKPKVKVQPKPKATKPKVVSQGRHWHVHALT-- 389
 DB 110 GPPRQHLHTPTPELS-----TGMLLRPVQ-----P-----SEGKRLLYKLSGG 153
 QY 390 RINIGLSPDEKYELDLHARVRNRGSGYQIADVGLKSGAGKTTLLTAAGSTLAQVRDRI 449
 DB 154 LINDGESPRARVYNNLVAQVVRPLRGSRVAFSLKGVGKTTIAATLGATFASIRGDRV 213
 QY 450 LAADADGAGNLADRVGQSGATADVLAEK-ELSHNDIRAHVSVAANLEVTAPABYS 508
 DB 214 VAVDANPDRTGLSKIPLETAATVROLHAGTIBRSVDVARYTSKPSGSEVLAASETP 273
 QY 509 SAQRALSDADMHFIADPASRPNVLVADCGAGFPDPLTRGVLTSVSGVVVAASVSDAQ 568
 DB 274 AVSEAPFADIVYRLIDLERFYGLVLDCEPGLSHWKSYLEKADALVVSSASIDGAR 333
 QY 569 QASVALDWLRNNGYQDLASRAQVYINIMPGSPNVAKYDLVRHFEQVQPPVVMFMDR 628
 DB 334 SASATLDMLDHGHEDLVRSINIAVINGRPRGKVMKVIDHFSRRC--AVQLVFDP 391
 QY 629 HIAAGTISIDLDPIYKRYLELAALSDFEBAQR 665
 DB 392 HLBEGAETIDRLRGRGTALTELAALVADGFGPAQR 428

RESULT 11

Q73VH1_MYCPA PRELIMINARY; PRT; 388 AA.
 ID Q73VH1_MYCPA PRELIMINARY; PRT; 388 AA.
 AC Q73VH1;
 DT 05-JUN-2004 (TrEMBLrel. 27, Created)
 DT 05-JUN-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUN-2004 (TrEMBLrel. 27, Last annotation update)
 DE Hypothetical protein.
 GN OrderedLocustNames=MAP3043C;
 OS Mycobacterium paratuberculosis.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacterineae; Mycobacteriaceae; Mycobacterium;
 OC Mycobacterium avium complex (MAC).
 NC NCB1_TaxID=1770;
 RN 11
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=K10;
 RA Li L., Bannantine J., Zhang Q., Amonsin A., Alt D., Kapur V.,
 RL Submitted (SRR-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB017238; AAS05551.1; -; Genomic_DNA.
 KW Complete proteome.
 SQ SEQUENCE 388 AA; 42040 MM; 60868AD8317FD09 CRC64;

Query Match 15.0%; Score 532.5; DB 2; Length 388;
 Best Local Similarity 34.2%; Pred. No. 1.5e-14;
 Matches 153; Conservative 54; Mismatches 162; Indels 79; Gaps 11;
 QY 229 GHRVTDTERNVGKATGPSIOARLRAEASGAOLAPGTEPSAPLQOPRSYLAPTTPRA 288

DB 6 GHKKYTGAMNRG--AIRGAVQOPAAAGAPGAPRPGMWRGNDP-----AAATVPQ 55
 QY 289 PTEP-----PSPSPQRSGRAERVRVADLAAQHAAPDSITATTGCR--RRKRAAP 342
 DB 56 PHGAPGRPRAPS-----QRVSP-----PARPTTGLATLDRPDQEDBAR 98
 QY 343 LDATQKSLRPAAKPKVKVQPKPKATKPKVVSQGRHWHVHALTRINIGLSPDEKYE 402
 DB 99 PD-----MGLVRLTGLIDGPKKATYE 122
 QY 403 LDHARVRNRGSGYQIADVGLKSGAGKTTLLTAAGSTLAQVRDRIADADGAGTAA 462
 DB 123 NELEAIRAAVGAFFIYAVNLKGVGKTVVEALGSTFAVRDRLALDID--AGDLA 180
 QY 463 DRVQSGATADVLAEKELSHYNDIRAHVSVAANLEVTAPAPSSAQRALSDADMHFI 522
 DB 181 ERHGRHNPMSMANILRGGPATRIEDIALTYMNGFRLEVLGDPYATIDRLEKQDVYKA 240
 QY 523 ADPASRPNVLVADCGAGFPDPLTRGVLTSVSGVVVAASVSDAQASVALDWLRNNGY 582
 DB 241 FSLMRNHYSVVLVDCVKAHNSVWEAVLPESBALVVTSPALDAIRKQTLLEWLRNNGY 300
 QY 583 QDLASRAQVYINIMPGSPN-VAVKDILVRHFEQVQGRVYVMFMDRHLAAGTISIDL 641
 DB 301 QRLMRSTYLVAVNHPEPAKVDAVATTELD--LSARVGAATVLPFDRHVEGRKIALDRL 357
 QY 642 DPIYKRYLELAALSDPF-----BRAGR 665
 DB 358 SKESRRSYLEMAALADMPGRGEQRGR 385

RESULT 12

08VKJ9_MYCTU PRELIMINARY; PRT; 405 AA.
 ID 08VKJ9_MYCTU PRELIMINARY; PRT; 405 AA.
 AC 08VKJ9;
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE Hypothetical protein.
 GN OrderedLocustNames=MT0552;
 OS Mycobacterium tuberculosis.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacterineae; Mycobacteriaceae; Mycobacterium;
 OC Mycobacterium tuberculosis complex.
 NC NCB1_TaxID=1773;
 RN 11
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=CDC 1551 / Oshkosh;
 RX MEDLINR=22206494; PubMed=12218036;
 RX DOI=10.1128/JB.184.19.5479-5490.2002;
 RA Fleischmann R.D., Alland D., Bisen J.A., Carpenter L., White O.,
 RA Peterson J.D., Deboy R.T., Dodson R.U., Gwinn M.L., Haft D.H.,
 RA Hickey B.K., Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D.,
 RA Salzberg S.L., Delcher A.L., Uterback T.R., Weidman J.F., Kouri H.M.,
 RA Gill J., Mikula A., Bishai W., Jacobs W.R. Jr., Venter J.C.,
 RA Fraser C.M.;
 RT "Whole-genome comparison of Mycobacterium tuberculosis clinical and
 RT laboratory strains.";
 RL J. Bacteriol. 184:5479-5490 (2002).
 DR EMBL; AB000516; AAK44775.1; -; Genomic_DNA.
 DR TIGR; MT0552; -;
 KW Hypothetical protein.
 SQ SEQUENCE 405 AA; 43093 MM; C62780825413B009 CRC64;

Query Match 14.9%; Score 530; DB 2; Length 405;
 Best Local Similarity 29.9%; Pred. No. 1.9e-14;
 Matches 155; Conservative 62; Mismatches 167; Indels 134; Gaps 14;
 QY 158 PGTPTGAPQOPEBPAPVSHGHQPRRTAPAPWAKPITGPPAPASPSASAPPT 214
 DB 7 PTGAGVAPDSGNGCTDH-----FTVQLPVPVSGVXPAAAGGT 45

```

Qy 215 PIRPAPQHSRRARRGKRYTDTDRBANGKATGTSIQARLPAEBSAQALPOTBSPAPL 274
    |||
    |||
Db 46 PIRSVVA-----GFRTO-----RLDPTAGAYTSGDEGFA-- 75
    |||
    |||
Qy 275 GQPRSVIAPTRBAPTEPP--SPSPQSRSGRAARRVHPDILAAQHAAQPSITAAT 331
    :|||
    :|||
Db 76 -----SPAREPPYRLPEP-----HTPYPEL----- 96
    :|||
    :|||
Qy 332 GGRRRKRAAPDIDATOKSLPRAAKGPKVKYKQKPKATPEPKVVSQGRHWHALT-- 388
    |||
    |||
Db 97 -----ATTTLRP-----VKP-----PP-----SEGRRLLYLLSGR 123
    :|||
    :|||
Qy 390 RINIGLSPBEKYEILDHARVRBNPRGSYOIAYVGLKGAGKTTLTALGSTIAQVRADI 448
    :|||
    :|||
Db 124 LINAEGPRAAHNDLVAOVNRLRGCTRIAYLSLKGAGVKTTITALTGATPADLRDRV 188
    :|||
    :|||
Qy 450 IALDADPGANLADRVGOSGATIADVLAEKE--LSHYNDIRAHTSVNAVNLVLPAPSEY 508
    :|||
    :|||
Db 184 VAVDANPDRGTLISQKVPLETPATVRHLRLDADGIERSDVRYGTSKSGSLAEVLASDP 243
    :|||
    :|||
Qy 508 SAQRALSDADWHPADPPASRFFNYLMLADCGAGFPDPLTREVLSVGSVVVAVSUIDAQ 566
    :|||
    :|||
Db 244 ASSDFASSADYRTTLDILERFYGLVLTDCGTGLHSMSAVALPRSDVLVAVSSSIDGAR 303
    :|||
    :|||
Qy 569 QASVALDMLRNNGYODLASRACVIVIHIMGESEPNVAVKDILVRHPEQOVQBPVVWEMWR 628
    :|||
    :|||
Db 304 SAAATITDWMQAGHGDQOVNSIIVAVNAVPRAGKVQGVKVEHPSRCRAVRVY--PDP 361
    :|||
    :|||
Qy 628 HTAAGTEISLDLIDLPYKRVLELAALSDFEBAQR 666
    :|||
    :|||
Db 362 HLEBGAEIALDLRLRETRREALTIEAALVAVAGFQEDPR 399
    :|||
    :|||

```

Query Match	14.9%	Score 530;	DB 2;	Length 405;
Best Local Similarity	29.9%	Pred. No. 1.9e-14;		
Matches 155;	Conservative 62;	Mismatches 167;	Indels 134;	Gaps 14;

158 PQTPGAQPPQSPAPAHVSHSGPHQPRRTAPAPPAKMPICGPPAPs---RPSASPAEP 214

[illegible]

	RESULT	14
OY	ID	Q7UIU4_MYCBO
Dt	PRT;	405 AA.
AC	Q7UIU4;	
Dt	01-OCT-2003	(TREMBLrel. 25, Created)
Dt	01-OCT-2003	(TREMBLrel. 25, Last sequence update)
Dt	01-OCT-2003	(TREMBLrel. 25, Last annotation update)
DE	Hypothetical protein MD0543.	
GN	OrderedLocustNames=MD0543;	
OS	Mycobacterium bovis.	
OC	Bacteria; Actinobacteria; Actinomycetales;	
OC	Corynebacterineae; Mycobacteriaceae; Mycobacterium;	
OX	NCBI_TaxID=1765;	
RN	[1]	
RP	NUCLEOTIDE SEQUENCE.	
RC	STRAIN=AF2122/97;	
RX	MEDLINE=22709107; PubMed=12786972; DOI=10.1073/pnas.1130426100;	
RA	Garnier T., Bighelmier K., Camus J.-C., Medina N., Mansoor H.,	
RA	Pyror M., Dufroy S., Gordin S., Lacroix C., Monsempé C., Simon S.,	
RA	Harries B., Atkin R., Doggett J., Mayes R., Keating I., Wheeler P.R.,	
RA	Parshall J., Barrell B.G., Cole S.T., Gordon S.V., Hewison R.G.;	
RT	"The complete genome sequence of Mycobacterium bovis."	
RL	Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882(2003).	
KW	EMBL; BX248335; CAD93405.1; -; Genomic_DNA.	
SV	Complete proteome; Hypothetical protein.	
SQ	SEQUENCE 405 AA; 43054 MW; 982503B09A1D52 CRC64;	
Query Match	14.9%; Score 530; DB 2; Length 405;	
Best Local Similarity	29.9%; Pred. No. 1.9e-14;	
Matches 155; Conservative	62; Mismatches 167; Indels 134; Gaps 14;	
Oy	158 FQTPTGAPQQESPAFHYVSHGHQPQPRTAEPAPMAKPIGCEPPAPAS--RPSASPAPP	214
Dd	7 PRITVGAGDSNGGGDH-----FTYGDLFPVPVSVAIPAAAGSET	45
Oy	215 FTTPAPQHSRRARRGRHRYTDERNVGVATGPSIQARLRAREKSAGAOIADGTETSPAPLV	274


```

Db      46 PTRSVA-----GRRTO-----RUDPTAYGAYSGPDGPA-- 75
Qy      275 GQPSYLAPTPRAPTEPPP---SPSPQNRSGRABRRVHPDLAQAHAAPDSITATY 331
Db      76 -----SPAERPPYRLBVP-----HTYPEL----- 96
Qy      332 GGRRRKAPADLDATOKSLRPAKGPVKYKVPQPKATKPKVVSQRGMRVWYHALT-- 389
Db      97 -----ATTTLARP-----VKP-----PSGWRRLYLISGR 123
Qy      390 RINIGLSPDEKYEYLDLHARVRNPRGSGYQIAVVGKGGAGKTTTLTAALGSLAQRDRI 449
Db      124 LINGEGPRAMHLDLVAAQVNRPLRGCTRIAVLSIKGGVGKTTTATLGAFTADLRGPRV 183
Qy      450 LALADPAGNLADRVGRGSGATTADVLABKG-LSHVNDIRAHSVNAVNLLEVLPAPRYS 508
Db      184 VAVNAPDRGTLTGKPLETATVRLHRLDADGIERSDVGRYTSKGPGLVLAISDSDP 243
Qy      509 SAQRALSDADWHFIADPAPSRFYNLVLADCGAGFPDPLTRGVLTSGVGVVAVSYSDQAQ 568
Db      244 ASSAPFSAADYTRTLDIRTYGLVLTDCGTGLHSMSAVLPRSDVLVVVSSGSDGAR 303
Qy      569 QASVALDMLRNGYQDLASRACVVINIIMPGEPRVAVDVLVRFHQVQVQGVVMPMDR 628
Db      304 SAAATLDWLQMHGDDQVNRGIAVNAVPRAGKVDVGKVEHFSRRCAVRVY--PPDP 361
Qy      629 HIAAGTISLDLPIYKRYLRLAALSDPBRAGR 666
Db      362 HLEGGAILDLRLRRTREALTILAAVVAAGPDPDR 399

```

RESULT 15

```

O6SSB8 CHLRE
ID O6SSB8_CHLRE PRELIMINARY; PRT; 3889 AA.
AC O6SSB8;
DT 05-JUL-2004 (TREMBlrel. 27, Created)
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
DE Minus agglutinin.
OS Name=SAD1;
ON Chlamydomonas reinhardtii.
OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae;
OC Chlamydomonadales; Chlamydomonadaceae; Chlamydomonas.
OX NCBI_TaxID=3055;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CC-621;
RX PubMed=15659633; DOI=10.1105/lpc.104.028035;
RA Perle P.J., Waffenschmidt S., Umen J.G., Lin H., Lee J.H., Ishida K.,
RA Kubo T., Lau J., Goodenough U.W.;
RT "Plus and Minus Sexual Agglutinins from Chlamydomonas reinhardtii.";
RL EMBL; AY450929; AAS07042.1; -; Genomic DNA.
DR GO; GO:0005139; F:structural constituent of cell wall; IEA.
DR InterPro; IPR002951; Atrophin.
DR InterPro; IPR006315; Autotransporter.
DR PRINTS; PRO1222; ATROPHIN.
DR PRINTS; PRO1218; PSTLXTENSIN.
DR TRIPFAMS; TIGR01414; autotrans_bar1; 1.
SQ SEQUENCE 3889 AA; 389223 MW; 4DB5B44D5507214A CRC64;

```

Query Match 13.9%; Score 495; DB 2; Length 3889;
 Best Local Similarity 33.3%; Pred. No. 4.4e-12;
 Matches 125; Conservative 35; Mismatches 143; Indels 72; Gaps 15;

```

Qy      17 APDMMAGPFPDPSPAPASANLPRKNGQTP---PRTSDLSKRVASAPPPPPPP 73
Db      925 SPSPSPAPSPPPSPSPSPAPSPSPSPSPAPDLPPSPSPSPSPSPAPSPSPSP 984
Qy      74 PPPTPMPIAGBPPSPSPAPSKPTPPMPIAGBPAPPKPTPPMPIAGBPAPPKPT 133

```

```

Db      985 PEPSPAPLP--PPSPSPSPAPSPSPPP--SPSPSPAPSPSPPP--SPSPSPAPPS 1038
Qy      134 PMPIAGPAPPTPESQLAP--PRPPTQPTTGAPQOQESDPAHVPSHGPHQPRATAPAP 191
Db      1039 PEPSPSP--PSAPSPAPSPSPSPSPAPSPSPSPAPSPSPSPAPSPSPSPAPSP 1097
Qy      192 WAKNPISGP---PRAPSRPSAPAP--PTRAPQHSRRARGRHRYTDTERNYGVATG 246
Db      1098 SPREPSPSPSPVPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSP 1131
Qy      247 PSIOARLRABASGAQLAPGTE--PSAPALQPRS-----YLAPTPRAPTEPPSPSQRN 301
Db      1132 PSLEP---PSAPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSP 1188
Qy      302 SGRABRRVHVDL---AAQHAQAQPDSTATTGRRRRKAPDLDATOKSLRPAKGP 358
Db      1189 A-----PPLPLPSPTQSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSP 1230
Qy      359 VKKVPQPKPKATKPP 373
Db      1231 PPSAPQAPSSPPPP 1245

```

Search completed: April 14, 2006, 17:32:20
 Job time : 250.732 secs

This Page Blank (uspto)

Db 181 HOPRRAPAPMAKMWIGBPPAPSPBSASPAEPTRPAQHSRRARRGHRYTDERNV 240
Qy 241 GKATAPSIQARABASGALAPGBSPAPLGGPRSYLAPPTPAPTPPSPSPOR 300
Db 241 GKATAPSIQARABASGALAPGBSPAPLGGPRSYLAPPTPAPTPPSPSPOR 300
Qy 301 NSGRABRRVHPDLAAQHAAPDSITTAATGGRRRRRAAPDLATQKSLRPAKGPVK 360
Db 301 NSGRABRRVHPDLAAQHAAPDSITTAATGGRRRRRAAPDLATQKSLRPAKGPVK 360
Qy 361 KYPQPKATKPKKVSQSGMRWVHALTRINGLSPDEKYEIDLARVRNRGSGYQA 420
Db 361 KYPQPKATKPKKVSQSGMRWVHALTRINGLSPDEKYEIDLARVRNRGSGYQA 420
Qy 421 VVGLKGGAGKTTTLAALGSTLAQVRADRIALADPAAGNLADRVGSGATTAADYLA 480
Db 421 VVGLKGGAGKTTTLAALGSTLAQVRADRIALADPAAGNLADRVGSGATTAADYLA 480
Qy 481 ELSHYNDIRAHTSVNAVLEVLPAPEYSSAQRALSDADWHFIADPASRFTNLVLADCG 540
Db 481 ELSHYNDIRAHTSVNAVLEVLPAPEYSSAQRALSDADWHFIADPASRFTNLVLADCG 540
Qy 541 FPDPLTRGVLTSTVSGVVVAVSVSIDGAQASVALDMLRNNGYODLASRACVVIHIMP 600
Db 541 FPDPLTRGVLTSTVSGVVVAVSVSIDGAQASVALDMLRNNGYODLASRACVVIHIMP 600
Qy 601 PNVAVKDLVHRFEQOVQGRVVMVPMDRHIAAGTEISLDLDPYKRKYLELAALSD 660
Db 601 PNVAVKDLVHRFEQOVQGRVVMVPMDRHIAAGTEISLDLDPYKRKYLELAALSD 660
Qy 661 ERAGRR 666
Db 661 ERAGRR 666

RESULT 2
US-09-050-739-94
; Sequence 94, Application US/09050739
; Patent No. 6641814
; GENERAL INFORMATION:
; APPLICANT: ANDERSEN, Peter
; APPLICANT: NIELSEN, Rikke
; APPLICANT: OTTINGER, Thomas
; APPLICANT: RASMUSSEN, Peter Birk
; APPLICANT: ROSENKRANDS, Ida
; APPLICANT: WELDINGH, Karin
; APPLICANT: FLORIO, Walter
; TITLE OF INVENTION: NUCLEIC ACIDS FRAGMENTS AND POLYPEPTIDE FRAGMENTS
; FILE REFERENCE: 670001-2002.1
; CURRENT APPLICATION NUMBER: US/09/050,739
; CURRENT FILING DATE: 1998-03-30
; EARLIER APPLICATION NUMBER: 0376/97
; EARLIER FILING DATE: 1997-04-02
; EARLIER APPLICATION NUMBER: 1277/97
; EARLIER FILING DATE: 1997-11-10
; EARLIER APPLICATION NUMBER: 60/044,624
; EARLIER FILING DATE: 1997-04-18
; EARLIER APPLICATION NUMBER: 60/070,488
; EARLIER FILING DATE: 1998-01-05
; NUMBER OF SEQ ID NOS: 173
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 94
; LENGTH: 308
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-09-050-739-94

Query Match 44.3%; Score 1575; DB 2; Length 308;
Best Local Similarity 99.7%; Pred. No. 6.1e-103;
Matches 307; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 359 VKKVKQPKATKPKKVSQSGMRWVHALTRINGLSPDEKYEIDLARVRNRGSGYQ 418
Db 1 MKKVKQPKATKPKKVSQSGMRWVHALTRINGLSPDEKYEIDLARVRNRGSGYQ 60
Qy 419 IAVVGLKGGAGKTTTLAALGSTLAQVRADRIALADPAAGNLADRVGSGATTAADYLA 478
Db 61 IAVVGLKGGAGKTTTLAALGSTLAQVRADRIALADPAAGNLADRVGSGATTAADYLA 120
Qy 479 EKELSHYNDIRAHTSVNAVLEVLPAPEYSSAQRALSDADWHFIADPASRFTNLVLADCG 538
Db 121 EKELSHYNDIRAHTSVNAVLEVLPAPEYSSAQRALSDADWHFIADPASRFTNLVLADCG 180
Qy 539 AGFPDLTRGVLTSTVSGVVVAVSVSIDGAQASVALDMLRNNGYODLASRACVVIHIMP 598
Db 181 AGFPDLTRGVLTSTVSGVVVAVSVSIDGAQASVALDMLRNNGYODLASRACVVIHIMP 240
Qy 599 GEPNVAVKDLVHRFEQOVQGRVVMVPMDRHIAAGTEISLDLDPYKRKYLELAALSD 658
Db 241 GEPNVAVKDLVHRFEQOVQGRVVMVPMDRHIAAGTEISLDLDPYKRKYLELAALSD 300
Qy 659 DEERAGRR 666
Db 301 DEERAGRR 308

RESULT 3
US-09-461-774-10
; Sequence 10, Application US/09461774
; Patent No. 6887481
; GENERAL INFORMATION:
; APPLICANT: CHAN, Lily
; APPLICANT: CHUNG, Maxey Ching Ming
; APPLICANT: LIM, Renee Lay Hong
; TITLE OF INVENTION: Bacterial-derived molecules and therapeutic and
; TITLE OF INVENTION: diagnostic uses thereof
; FILE REFERENCE: 1781-0180P
; CURRENT APPLICATION NUMBER: US/09/461,774
; CURRENT FILING DATE: 1999-12-15
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 10
; LENGTH: 386
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-09-461-774-10

Query Match 11.6%; Score 410.5; DB 2; Length 386;

Best Local Similarity 33.3%; Pred. No. 4.2e-21; Indels 79; Gaps 12;

Matches 109; Conservative 20; Mismatches 119; Indels 79; Gaps 12;

Qy 15 MEAPD-----DMAQPFDPDSASFPAPASANTLKPNGQTPPTSDDLSERFVSA 64
Db 66 LEAPNPVPPAPPGVSAAPAPMPPTPLPVPSPGSAARVPVAVP-----M 113
Qy 65 PP-----PPPPPPPPPT-----PPIAAGBPSEPPASAKPPTPMTIAGE-PA 110
Db 114 PAPKMPALPAPAPAPSPPTSWLAVPVPVPLVVKMPPSPVPPPPABETPN 173
Qy 111 PKKPTPPPIA-GEPAPPKPPTPM-----PIAGPATPTESQIAPRPTPTPTNG 163
Db 174 PAPAPVLENSPPPPPPVPPVPPVPLTNPVPPAPPAANTSNLSLRPAPAPPLK 233
Qy 164 APOQSPAPVHVSHPQPRRTAPAPWAKMPICGPPAPAPSPASPAEPTTPAPQHS 223
Db 234 PAPAPVLENSPPPPPPVPPVPPVPLTNPVPPAPPAANTSNLSLRPAPAPPLK 286
Qy 224 RARRGHRYTDERNVGKATGPGTQA---RLRAEAS--GAQLAGTBSAPALG 277
Db 287 -----LPGSPAPAPAPPLPNSPAPAPGAPAPAPAPAPAPAPAPAPAP 323
Qy 278 RSYLAPPTPAPTPPSPSPQNSGR 304
Db 324 LPYSPAPAPCPVPGAPLAPLPIGR 350

```
RESULT 4
US-08-642-255-32
; Sequence 32, Application US/08642255
; Patent No. 5773249
; GENERAL INFORMATION:
; APPLICANT: CAPPELLO, Joseph
; APPLICANT: FERRARI, Franco A.
; TITLE OF INVENTION: High Molecular Weight Collagen-Like
; TITLE OF INVENTION: Protein Polymers
; NUMBER OF SEQUENCES: 135
; CORRESPONDENCE ADDRESS:
; ADDRESSER: FLEHR, HOBACH, TEST, ALBRITTON & HERBERT
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/642,255
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: ROWLAND, Bertram I.
; REGISTRATION NUMBER: 20,015
; REFERENCE/DOCKET NUMBER: A55556-3/BIR
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 494-8700
; TELEFAX: (415) 494-8771
; TELEX: 910 277299 FHT UR
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 330 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULAR TYPE: protein
; US-08-642-255-32

Query Match          10.2%; Score 364; DB 1; Length 330;
Best Local Similarity 33.8%; Pred. No. 6,4e-18;
Matches 101; Conservative 8; Mismatches 112; Indels 78; Gaps 15;

QY 29 PSASPPAPA-----SANTLPKNGQTPPTSDLSKRVSAAPPPPPPPPPPTPM-PIA 83
DB 57 PCPPGPPGAGPVGSPAPRPPRPPGPP-----GAPGPPGPPGPPGPPAGPV- 108
QY 84 AGSPSPBPASAKPPTPMPDIAGP-BDAPPKPPTPMPDIAGP-----BPAPKPTPMP 137
DB 109 -GSPGAPGPPG--PCGPPGPPGAPGPPGPPGPPGPPGPP-AGPVGSPGAPGPPGPPGPP 164
QY 138 IAGPAPPTTSQALAPRPPPTQTPTGAPQPPSPAPRVNPSHGHPKRTA-----PAPPA 193
DB 165 PGAPGPP--GPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPP 221
QY 194 KMPIGP-----PPASRPSASPAE--PPTBPAPQHSRARAGHRYTDTSENVKVA 244
DB 222 AGPVGSPGAPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPP 257
QY 245 TGPSIGARLALAEASGAQLAGTSPAPVAGQPRSYLAPPTRAPPTPPSPSPQRNSG 303
DB 258 PGP-----AGPVGSPGAPGPPGPPGPP-----GPGAPGPPGPPGPPGPPGPPGPPAG 301

RESULT 5
US-07-609-716-65
; Sequence 65, Application US/07609716
```

```
; Patent No. 5514581
; GENERAL INFORMATION:
; APPLICANT: Ferrari, Franco A.
; APPLICANT: Cappello, Joseph
; TITLE OF INVENTION: Functional Recombinantly Prepared
; TITLE OF INVENTION: Synthetic Protein Polymer
; NUMBER OF SEQUENCES: 118
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Flehr, Hobach, Test, Albritton & Herbert
; STREET: Four Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: CA
; COUNTRY: US
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/609,716
; FILING DATE: 06-NOV-1990
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Rowland, Bertram I.
; REGISTRATION NUMBER: 20015
; REFERENCE/DOCKET NUMBER: A-55186-3/BIR
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-781-1989
; TELEFAX: 415-398-3249
; INFORMATION FOR SEQ ID NO: 65:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 408 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULAR TYPE: protein
; US-07-609-716-65

Query Match          10.2%; Score 364; DB 1; Length 408;
Best Local Similarity 33.8%; Pred. No. 8,2e-18;
Matches 101; Conservative 8; Mismatches 112; Indels 78; Gaps 15;

QY 29 PSASPPAPA-----SANTLPKNGQTPPTSDLSKRVSAAPPPPPPPPPPTPM-PIA 83
DB 57 PCPPGPPGAGPVGSPAPRPPRPPGPP-----GAPGPPGPPGPPGPPAGPV- 108
QY 84 AGSPSPBPASAKPPTPMPDIAGP-BDAPPKPPTPMPDIAGP-----BPAPKPTPMP 137
DB 109 -GSPGAPGPPG--PCGPPGPPGAPGPPGPPGPPGPPGPP-AGPVGSPGAPGPPGPPGPP 164
QY 138 IAGPAPPTTSQALAPRPPPTQTPTGAPQPPSPAPRVNPSHGHPKRTA-----PAPPA 193
DB 165 PGAPGPP--GPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPP 221
QY 194 KMPIGP-----PPASRPSASPAE--PPTBPAPQHSRARAGHRYTDTSENVKVA 244
DB 222 AGPVGSPGAPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPP 257
QY 245 TGPSIGARLALAEASGAQLAGTSPAPVAGQPRSYLAPPTRAPPTPPSPSPQRNSG 303
DB 258 PGP-----AGPVGSPGAPGPPGPPGPP-----GPGAPGPPGPPGPPGPPGPPGPPAG 301

RESULT 6
US-08-475-411A-65
; Sequence 65, Application US/08475411A
; Patent No. 6140072
; GENERAL INFORMATION:
; APPLICANT: Ferrari, Franco A.
; APPLICANT: Cappello, Joseph
; TITLE OF INVENTION: Functional Recombinantly Prepared
; TITLE OF INVENTION: Synthetic Protein Polymer
```


Db 258 RGP-----AGVSGPAGFPGPGRPP-----GPRGAGPGRPGPGRPGAG 301

RESULT 8

US-07-757-022B-14

Sequence 14, Application US/07757022B

Patent No. 6433142

GENERAL INFORMATION:

APPLICANT: Gesner, Thomas G.

APPLICANT: Clark, Stephen C.

APPLICANT: Turner, Katherine

APPLICANT: Hewick, Rodney M.

TITLE OF INVENTION: Megakaryocyte Stimulating Factors

NUMBER OF SEQUENCES: 143

CORRESPONDENCE ADDRESS:

ADDRESSER: Genetics Institute, Inc.

STREET: 87 Cambridgepark Drive

CITY: Cambridge

STATE: Massachusetts

COUNTRY: U.S.A.

ZIP: 02140

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/757,022B

FILING DATE: 19910910

CLASSIFICATION: 530

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/643,502

FILING DATE: 18-JAN-1991

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/546,114

FILING DATE: 29-JUN-1990

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/457,196

FILING DATE: 29-DEC-1989

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/390,901

FILING DATE: 08-AUG-1989

ATTORNEY/AGENT INFORMATION:

NAME: Gesner, Luann

REGISTRATION NUMBER: 31,822

REFERENCE/DOCKET NUMBER: GI 5190

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617)876-1170

TELEFAX: (617)876-5851

INFORMATION FOR SEQ ID NO: 14:

SEQUENCE CHARACTERISTICS:

LENGTH: 941 amino acids

TYPE: AMINO ACID

TOPOLOGY: linear

MOLECULAR TYPE: protein

US-07-757-022B-14

Query Match 9.8%; Score 347; DB 2; Length 941;

Best Local Similarity 25.4%; Pred. No. 3.4e-16;

Matches 146; Conservative 58; Mismatches 259; Indels 112; Gaps 29;

QY 28 DPASPP--PAPASANLPKNGOPPTSDLSERFVSAAPPPTP-----PPPEPTPM 80

Db 251 BPAPTPPKERTPTPKBPATTKBPATTP--KEPAPTAKKBPATTPPKBPATTPKEPA 308

QY 81 PIAGBP-----PSPBPASKPPTP--PMDIAGPEPAP--PKP 114

Db 309 PTTTKBSPTTPKBPATTKSAPTTKBPATTKSAPTTKKPSPTTKKBPATTPKE 368

QY 115 PTP-----PMDIAGPEPAP--PKPPTP-----PMDI--GPAP--TPTESQAPRPPTPO 159

Db 369 PAPTTPKBPATTPKBPATTPKBPATTKKBPATTPKBPATTPKBPATTPKBPATTPK 428

QY 160 TPTG-AFOQESPAHPVSH-----GPHQPRITAPAPWAMPICGEPAPASRS-ASPAL 213

Db 429 TPETLAPTTPEKAPATTPSEIAPTTPEEPPTTPBEPATTPPKAAANTPKBPATTPKE 488

QY 214 P-----PTRPAPQHSRRARRHRRTDERVGVATGSPSIOANRLABEASGAOLAPGT-- 267

Db 489 PAPTTPKBPATTPK-----ETAPTPKGAATTLKBPATTPKBPATTPKBPATTPKE 540

QY 268 EPS-----PAPIGQPSRYLAPPTRPAPTEP--PSPSPQNSGRARRVHVDIAQHA 320

Db 541 EPSTTSKAPATTPPKGTATTPKBPATTPKBPATTPKBPATTPKBPATTPKBPATTPK 599

QY 321 AQPDSITATGGRRRRGAAPDIDATOKSLRPAKPK-VKQVQKPKAT--KPKV 376

Db 600 --PKELAPTTKGTSTTSKBPATTPKBPATTP--PKBPATTPKBPATTPPETPT 655

QY 377 SQGMRHVAALRINIGLSPDEKYLIDLAR-----VRBNPGSYOIAVGLKGGAGKT 431

Db 656 SEVSTPTTKPPTTH--KSPDSTP-ELSAEPKALNSPK-----BPVPTT 702

QY 432 TLTAAGSTLAQVRADRIALD--ADGAGNLADRVGSGGATLADVLAKELSHVNDIR 489

Db 703 KTPAATPEKNTTAKDKITTRDARTTPTTAPAKPKTKETATTEKTESKITATTTQVT 762

QY 490 AHTSVANLVLP-----APRYSSAQRALSDAD 518

Db 763 STTQDTPPKITTLKTTLAPKVTTKKITTTTE 797

RESULT 9

US-07-757-022B-84

Sequence 84, Application US/07757022B

Patent No. 6433142

GENERAL INFORMATION:

APPLICANT: Gesner, Thomas G.

APPLICANT: Clark, Stephen C.

APPLICANT: Turner, Katherine

APPLICANT: Hewick, Rodney M.

TITLE OF INVENTION: Megakaryocyte Stimulating Factors

NUMBER OF SEQUENCES: 143

CORRESPONDENCE ADDRESS:

ADDRESSER: Genetics Institute, Inc.

STREET: 87 Cambridgepark Drive

CITY: Cambridge

STATE: Massachusetts

COUNTRY: U.S.A.

ZIP: 02140

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/757,022B

FILING DATE: 19910910

CLASSIFICATION: 530

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/643,502

FILING DATE: 18-JAN-1991

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/546,114

FILING DATE: 29-JUN-1990

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/457,196

FILING DATE: 29-DEC-1989

ATTORNEY/AGENT INFORMATION:

NAME: Gesner, Luann

REGISTRATION NUMBER: 31,822

REFERENCE/DOCKET NUMBER: GI 5190

TELECOMMUNICATION INFORMATION:

Db 768 KTPATKPEMTTAKDKTTERDLTTPETTTAAKMTKEATTTKTESKITATTQVT 827
Qy 490 AHTSVANVLEVL-----APEYSSAQRALSDAD 518
Db 828 STTQDTPPKITTLKTTTLAPKVTTKTITTTTE 862

RESULT 11

US-07-757-022B-58
; Sequence 58, Application US/07757022B
; Patent No. 6433142
; GENERAL INFORMATION:
; APPLICANT: Gesner, Thomas G.
; APPLICANT: Clark, Stephen C.
; APPLICANT: Turner, Katherine
; APPLICANT: Hewick, Rodney M.
; TITLE OF INVENTION: Megakaryocyte Stimulating Factors
; NUMBER OF SEQUENCES: 143
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 Cambridgepark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/757,022B
; FILING DATE: 19910910
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/643,502
; FILING DATE: 18-JAN-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/546,114
; FILING DATE: 29-JUN-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/457,196
; FILING DATE: 29-DEC-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/390,901
; FILING DATE: 08-AUG-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Cseert, Luann
; REGISTRATION NUMBER: 31,822
; REFERENCE/DOCKET NUMBER: GI 5190
; TELEPHONE: (617)876-1170
; TELEFAX: (617)876-5851
; INFORMATION FOR SEQ ID NO: 58:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1049 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-757-022B-58

Query Match 9.8%; Score 347; DB 2; Length 1049;

Best Local Similarity 25.4%; Pred. No. 3,9e-16;

Matches 146; Conservative 58; Mismatches 259; Indels 112; Gaps 29;

Qy 28 DPSASFP--PAPASANLPKENGQTPPTSDLSRFVSAPPPTP-----PPPPPTPTM 80
Db 359 BPAVTPKBPPTTPKBPATTKBPATTP--KEPAPTAPEKBPATTPKBPATTPKBP 416
Qy 81 PIAAGEP-----PSPEPAASKPP--PMIAGEBP--PKP 114
Db 417 PTTKEPSPTTPKBPATTKSAPTTPKBPATTKKBPATTKKBPATTKKBPATTKKBP 476

Qy 115 PTP-----PMPIAGEBPAP--PKPPTP-----PMPIA--GPAP--TPTESOLAPRPPTPQ 159
Db 477 PAPTPPKBPAPTPPKBPATTPKBPATTPKBPATTPKBPATTPKBPATTPKBPATTP 536
Qy 160 TPNG-APOQESPAHPHSH-----GPHQPRRTAPAPWAMGPIGEPAPASRPS--ASPAP 213
Db 537 TPBKLAFTPEKBPATTPBBLAFTTPBEPPTTPBEPAPTPPKAANTTPKBPATTPK 596
Qy 214 P-----PTPAPQHSRRARGRHYRTDTERVGVKATGPSIOARLASEASGAQAPGT-- 267
Db 597 PAPTPKBPATTPK-----ETAPTTKGTAPTLKBPAPTPPKBPATTPKBPATTPK 648
Qy 268 EPS-----PAPLQPSRYLAAPTTPAPTPP--PPSPQNSGRARRVHPDLAQAHA 320
Db 649 EPTSTSDKBPATTPKGTAPTPKBPATTPKBPATTPKBPATTPKBPATTPKBPATTP 707
Qy 321 AQPDSITAAITGRRRRRAAPDLDATQKSLAPKAPK--VKVVPQPKAT--KPPRVV 376
Db 708 --PEKLAFTYTKGPTSTSDKBPATTPKETAFTT--BKEPAPTPPKBPATTPPTPT 763
Qy 377 SORGMHVAALFRINGLSPDEKYEIDLHAR-----VRNPRGSYOIAVGLKGAGKT 431
Db 764 SEVSTPTTKPTTIH--KSPDESTP--ELSAEPTPKALENSPK-----EGVPTT 810
Qy 432 TLTALGSTLAQVRAADRILALD--ADPGAGNLADRVGSGATTADVLAKELSHYNDIR 489
Db 811 KTPATKPEMTTAKDKTTERDLTTPETTTAAKMTKEATTTKTESKITATTQVT 870
Qy 490 AHTSVANVLEVL-----APEYSSAQRALSDAD 518
Db 871 STTQDTPPKITTLKTTTLAPKVTTKTITTTTE 905

RESULT 12

US-07-757-022B-104

; Sequence 104, Application US/07757022B

; Patent No. 6433142

; GENERAL INFORMATION:

; APPLICANT: Gesner, Thomas G.

; APPLICANT: Clark, Stephen C.

; APPLICANT: Turner, Katherine

; APPLICANT: Hewick, Rodney M.

; TITLE OF INVENTION: Megakaryocyte Stimulating Factors

; NUMBER OF SEQUENCES: 143

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Genetics Institute, Inc.

; STREET: 87 Cambridgepark Drive

; CITY: Cambridge

; STATE: Massachusetts

; COUNTRY: U.S.A.

; ZIP: 02140

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; OPERATING SYSTEM: IBM PC compatible

; SOFTWARE: Patentin Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/07/757,022B

; FILING DATE: 19910910

; CLASSIFICATION: 530

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/643,502

; FILING DATE: 18-JAN-1991

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/546,114

; FILING DATE: 29-JUN-1990

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/457,196

; FILING DATE: 29-DEC-1989

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/390,901

; FILING DATE: 08-AUG-1989


```
QY 181 HOPRRATAPAPMAKMEIGBPBPSPBSASPAEPPTPAPPOHSRRARRGRHYTTERNV 240
DB 181 HOPRRATAPAPMAKMEIGBPBPSPBSASPAEPPTPAPPOHSRRARRGRHYTTERNV 240
QY 241 GKVATGSPSTQARLRABEASGAOLAPGTEPSPAPIGQPSRYLAPPTPAPTEPPSPSPOR 300
DB 241 GKVATGSPSTQARLRABEASGAOLAPGTEPSPAPIGQPSRYLAPPTPAPTEPPSPSPOR 300
QY 301 NSGRABERRVHPDLAAOHAAPDPSITTAATTGRRRKRRAAPDLATOKSLRPAKGPKYK 360
DB 301 NSGRABERRVHPDLAAOHAAPDPSITTAATTGRRRKRRAAPDLATOKSLRPAKGPKYK 360
QY 361 KVPQKPKATKPKPKVVSQSGRMHWALTRINLGLSPDEKXELDLARVRNPRGSGYQIA 420
DB 361 KVPQKPKATKPKPKVVSQSGRMHWALTRINLGLSPDEKXELDLARVRNPRGSGYQIA 420
QY 421 VVGLKGAAGKTTTLTAALGSTLAQVRADRIALADADGAGNLADRVGROSGATTADYLAEK 480
DB 421 VVGLKGAAGKTTTLTAALGSTLAQVRADRIALADADGAGNLADRVGROSGATTADYLAEK 480
QY 481 ELSHYNDIRAHTSVNAVNLVLPAPPEYSSAQRALSDADWHFIADPASRFYNLVLADCGAG 540
DB 481 ELSHYNDIRAHTSVNAVNLVLPAPPEYSSAQRALSDADWHFIADPASRFYNLVLADCGAG 540
QY 541 FPDPLTRGYLSTVSGVVVAVSVSIDGAQOASVALDMLRNNGYODLASRAACVNIHIMPGE 600
DB 541 FPDPLTRGYLSTVSGVVVAVSVSIDGAQOASVALDMLRNNGYODLASRAACVNIHIMPGE 600
QY 601 PNVAVKDLVRHFEQOVQPGRVVVMPPMDRHAAGTEISLDLDPITYRKVLELAALSDPF 660
DB 601 PNVAVKDLVRHFEQOVQPGRVVVMPPMDRHAAGTEISLDLDPITYRKVLELAALSDPF 660
QY 661 ERAGRR 666
DB 661 ERAGRR 666
```

```
RESULT 2
US-09-804-980-70
; Sequence 70, Application US/09804980
; Publication No. US20030147897A1
; GENERAL INFORMATION:
; APPLICANT: Scatens Serum Institut
; APPLICANT: Anderson, Peter
; TITLE OF INVENTION: M. Tuberculosis Antigens
; FILE REFERENCE: 670001-2002.4
; CURRENT APPLICATION NUMBER: US/09/804,980
; NUMBER OF SEQ ID NOS: 257
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 70
; LENGTH: 666
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-09-804-980-70
```

```
Query Match 100.0%; Score 3552; DB 3; Length 666;
Best Local Similarity 100.0%; Pred. No. 2.5e-180;
Matches 666; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MAADYDKLFRPHGMEAPDDMAAOPFDDPSASFPAPASANTLPKPGQCTPPPTSDDLSEK 60
DB 1 MAADYDKLFRPHGMEAPDDMAAOPFDDPSASFPAPASANTLPKPGQCTPPPTSDDLSEK 60
QY 61 FVSAPPPPPPPPPPTPMPPIAAGBPSPBPASAKPPTPMPPIAGBPBPAPKPPPTPMP 120
DB 61 FVSAPPPPPPPPPPTPMPPIAAGBPSPBPASAKPPTPMPPIAGBPBPAPKPPPTPMP 120
QY 121 IAGBPAPKPPPTPMPPIAGBPPTPESQLAAPRPPTPTGTGAPOQPSAPAHVPSHG 180
DB 121 IAGBPAPKPPPTPMPPIAGBPPTPESQLAAPRPPTPTGTGAPOQPSAPAHVPSHG 180
```

```
QY 181 HOPRRATAPAPMAKMEIGBPBPSPBSASPAEPPTPAPPOHSRRARRGRHYTTERNV 240
DB 181 HOPRRATAPAPMAKMEIGBPBPSPBSASPAEPPTPAPPOHSRRARRGRHYTTERNV 240
QY 241 GKVATGSPSTQARLRABEASGAOLAPGTEPSPAPIGQPSRYLAPPTPAPTEPPSPSPOR 300
DB 241 GKVATGSPSTQARLRABEASGAOLAPGTEPSPAPIGQPSRYLAPPTPAPTEPPSPSPOR 300
QY 301 NSGRABERRVHPDLAAOHAAPDPSITTAATTGRRRKRRAAPDLATOKSLRPAKGPKYK 360
DB 301 NSGRABERRVHPDLAAOHAAPDPSITTAATTGRRRKRRAAPDLATOKSLRPAKGPKYK 360
QY 361 KVPQKPKATKPKPKVVSQSGRMHWALTRINLGLSPDEKXELDLARVRNPRGSGYQIA 420
DB 361 KVPQKPKATKPKPKVVSQSGRMHWALTRINLGLSPDEKXELDLARVRNPRGSGYQIA 420
QY 421 VVGLKGAAGKTTTLTAALGSTLAQVRADRIALADADGAGNLADRVGROSGATTADYLAEK 480
DB 421 VVGLKGAAGKTTTLTAALGSTLAQVRADRIALADADGAGNLADRVGROSGATTADYLAEK 480
QY 481 ELSHYNDIRAHTSVNAVNLVLPAPPEYSSAQRALSDADWHFIADPASRFYNLVLADCGAG 540
DB 481 ELSHYNDIRAHTSVNAVNLVLPAPPEYSSAQRALSDADWHFIADPASRFYNLVLADCGAG 540
QY 541 FPDPLTRGYLSTVSGVVVAVSVSIDGAQOASVALDMLRNNGYODLASRAACVNIHIMPGE 600
DB 541 FPDPLTRGYLSTVSGVVVAVSVSIDGAQOASVALDMLRNNGYODLASRAACVNIHIMPGE 600
QY 601 PNVAVKDLVRHFEQOVQPGRVVVMPPMDRHAAGTEISLDLDPITYRKVLELAALSDPF 660
DB 601 PNVAVKDLVRHFEQOVQPGRVVVMPPMDRHAAGTEISLDLDPITYRKVLELAALSDPF 660
QY 661 ERAGRR 666
DB 661 ERAGRR 666
```

```
RESULT 3
US-10-080-170-640
; Sequence 640, Application US/10080170
; Publication No. US20030129601A1
; GENERAL INFORMATION:
; APPLICANT: CODE, S.T.
; TITLE OF INVENTION: COMPARATIVE MYCOBACTERIAL GENOMICS AS A TOOL FOR
; IDENTIFYING TARGETS FOR THE DIAGNOSIS, PROPHYLAXIS OR
; TREATMENT OF MYCOBACTERIOSES
; FILE REFERENCE: 03495.0219
; CURRENT APPLICATION NUMBER: US/10/080,170
; PRIOR FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: 60/270,123
; NUMBER OF SEQ ID NOS: 652
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 640
; LENGTH: 666
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-10-080-170-640
```

```
Query Match 100.0%; Score 3552; DB 4; Length 666;
Best Local Similarity 100.0%; Pred. No. 2.5e-180;
Matches 666; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MAADYDKLFRPHGMEAPDDMAAOPFDDPSASFPAPASANTLPKPGQCTPPPTSDDLSEK 60
DB 1 MAADYDKLFRPHGMEAPDDMAAOPFDDPSASFPAPASANTLPKPGQCTPPPTSDDLSEK 60
QY 61 FVSAPPPPPPPPPPTPMPPIAAGBPSPBPASAKPPTPMPPIAGBPBPAPKPPPTPMP 120
DB 61 FVSAPPPPPPPPPPTPMPPIAAGBPSPBPASAKPPTPMPPIAGBPBPAPKPPPTPMP 120
QY 121 IAGBPAPKPPPTPMPPIAGBPPTPESQLAAPRPPTPTGTGAPOQPSAPAHVPSHG 180
DB 121 IAGBPAPKPPPTPMPPIAGBPPTPESQLAAPRPPTPTGTGAPOQPSAPAHVPSHG 180
```



```
/ GENERAL INFORMATION:
/ APPLICANT: ANDERSEN, Peter
/ APPLICANT: NIELSEN, Rikke
/ APPLICANT: OETTINGER, Thomas
/ APPLICANT: RASMUSSEN, Peter Birk
/ APPLICANT: ROSENKRANDS, Ida
/ APPLICANT: WELDLINGH, Karin
/ APPLICANT: FLORIO, Walter
/ TITLE OF INVENTION: NUCLEIC ACIDS FRAGMENTS AND POLYPEPTIDE FRAGMENTS
/ TITLE OF INVENTION: DERIVED FROM M. TUBERCULOSIS
/ FILE REFERENCE: 670001-2002.1A
/ CURRENT APPLICATION NUMBER: US/10/620,246
/ CURRENT FILING DATE: 2003-07-15
/ PRIOR APPLICATION NUMBER: 09/050,739
/ PRIOR FILING DATE: 1998-03-30
/ PRIOR APPLICATION NUMBER: 0376/97
/ PRIOR FILING DATE: 1997-04-02
/ PRIOR APPLICATION NUMBER: 1277/97
/ PRIOR FILING DATE: 1997-11-10
/ PRIOR APPLICATION NUMBER: 60/044,624
/ PRIOR FILING DATE: 1997-04-18
/ PRIOR APPLICATION NUMBER: 60/070,488
/ PRIOR FILING DATE: 1998-01-05
/ PRIOR APPLICATION NUMBER: 10/138,473
/ PRIOR FILING DATE: 2002-05-02
/ PRIOR APPLICATION NUMBER: 09/791,171
/ PRIOR FILING DATE: 2001-02-20
/ PRIOR APPLICATION NUMBER: 09/415,884
/ PRIOR FILING DATE: 1999-10-08
/ PRIOR APPLICATION NUMBER: 60/116,673
/ PRIOR FILING DATE: 1999-01-21
/ PRIOR APPLICATION NUMBER: 1281/98
/ PRIOR FILING DATE: 1998-10-08
/ NUMBER OF SEQ ID NOS: 173
/ SOFTWARE: Patentin Ver. 2.0
/ SEQ ID NO 70
/ LENGTH: 666
/ TYPE: PRF
/ ORGANISM: Mycobacterium tuberculosis
US-10-620-246-70

Query Match      100.0%; Score 3552; DB 4; Length 666;
Best Local Similarity 100.0%; Pred. No. 2.5e-180;
Matches 666; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  MAADYKLFRRPHGMEAPDDMAAOPFPDSASFPAPASANTPKRNGQTPPTSDLSRR 60
DB      1  MAADYKLFRRPHGMEAPDDMAAOPFPDSASFPAPASANTPKRNGQTPPTSDLSRR 60
QY      61  FVSAPPPPPPPPPPTPMPPIAAGSPSPBASAPPTPMPPIAGSPAPAPKPTPMP 120
DB      61  FVSAPPPPPPPPPPTPMPPIAAGSPSPBASAPPTPMPPIAGSPAPAPKPTPMP 120
QY      121  IAGSPAPAPKPTPMPPIAGSPAPPTPTESQIAPRPPTPTGAPQOQSPAPAHVSHGP 180
DB      121  IAGSPAPAPKPTPMPPIAGSPAPPTPTESQIAPRPPTPTGAPQOQSPAPAHVSHGP 180
QY      121  IAGSPAPAPKPTPMPPIAGSPAPPTPTESQIAPRPPTPTGAPQOQSPAPAHVSHGP 180
DB      121  IAGSPAPAPKPTPMPPIAGSPAPPTPTESQIAPRPPTPTGAPQOQSPAPAHVSHGP 180
QY      181  HOPRRRTAPAPPAKMPIGSPAPSPASAPPTPMPPIAAGSPAPAPKPTPMPPIAGSPAP 240
DB      181  HOPRRRTAPAPPAKMPIGSPAPSPASAPPTPMPPIAAGSPAPAPKPTPMPPIAGSPAP 240
QY      241  GKVATGPSIOARLRABEASGAQLAPGTEPSAPLQGPSTIAPPTPAPTEPPSPSPOR 300
DB      241  GKVATGPSIOARLRABEASGAQLAPGTEPSAPLQGPSTIAPPTPAPTEPPSPSPOR 300
QY      301  NSGRARERVRHDLAQAHAAPDSITTAATGGRRRKRAAPDLIDATQKSLRPAKGPVK 360
DB      301  NSGRARERVRHDLAQAHAAPDSITTAATGGRRRKRAAPDLIDATQKSLRPAKGPVK 360
QY      361  KYKPKKATKPKPKVVSQSGMRHWALTRINIGLSPDEKYTELDLHARRRNRGSGYQA 420
DB      361  KYKPKKATKPKPKVVSQSGMRHWALTRINIGLSPDEKYTELDLHARRRNRGSGYQA 420
```

```
QY      421  VVGLKGAGKTTTLTAALSGTLAQVRADRLIADDPGAGNLDNRVROSGATITADYLAK 480
DB      421  VVGLKGAGKTTTLTAALSGTLAQVRADRLIADDPGAGNLDNRVROSGATITADYLAK 480
QY      481  ELSHYNDIRAHVSVAVNLVETLPAPBYSAGORALSDADWHFIADPARRRYNLVLADCGG 540
DB      481  ELSHYNDIRAHVSVAVNLVETLPAPBYSAGORALSDADWHFIADPARRRYNLVLADCGG 540
QY      541  FPDPLTRGLSTVSGVVAVSIDGAQASVALDMLRNGYQDLASRACVVNIHMPGE 600
DB      541  FPDPLTRGLSTVSGVVAVSIDGAQASVALDMLRNGYQDLASRACVVNIHMPGE 600
QY      601  PVAAYVDLVRHFEQVQVGRVVMWMDRIIAAGTEISLDLDPYKRTLELAALSDPF 660
DB      601  PVAAYVDLVRHFEQVQVGRVVMWMDRIIAAGTEISLDLDPYKRTLELAALSDPF 660
QY      661  ERAGRR 666
DB      661  ERAGRR 666
```

RESULT 6

```
US-10-080-170-640
/ Sequence 640, Application US/10080170
/ Publication No. US20040121322A9
/ GENERAL INFORMATION:
/ APPLICANT: COLE, S.T.
/ TITLE OF INVENTION: COMPARATIVE MYCOBACTERIAL GENOMICS AS A TOOL FOR
/ TITLE OF INVENTION: IDENTIFYING TARGETS FOR THE DIAGNOSIS, PROPHYLAXIS OR
/ TITLE OF INVENTION: TREATMENT OF MYCOBACTERIOSES
/ FILE REFERENCE: 03495.0218
/ CURRENT APPLICATION NUMBER: US/10/080,170
/ CURRENT FILING DATE: 2002-06-10
/ PRIOR APPLICATION NUMBER: 60/270,123
/ PRIOR FILING DATE: 2001-02-22
/ NUMBER OF SEQ ID NOS: 652
/ SOFTWARE: Patentin Ver. 2.1
/ SEQ ID NO 640
/ LENGTH: 666
/ TYPE: PRF
/ ORGANISM: Mycobacterium tuberculosis
US-10-080-170-640
```

```
Query Match      100.0%; Score 3552; DB 4; Length 666;
Best Local Similarity 100.0%; Pred. No. 2.5e-180;
Matches 666; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  MAADYKLFRRPHGMEAPDDMAAOPFPDSASFPAPASANTPKRNGQTPPTSDLSRR 60
DB      1  MAADYKLFRRPHGMEAPDDMAAOPFPDSASFPAPASANTPKRNGQTPPTSDLSRR 60
QY      61  FVSAPPPPPPPPPPTPMPPIAAGSPSPBASAPPTPMPPIAGSPAPAPKPTPMP 120
DB      61  FVSAPPPPPPPPPPTPMPPIAAGSPSPBASAPPTPMPPIAGSPAPAPKPTPMP 120
QY      121  IAGSPAPAPKPTPMPPIAGSPAPPTPTESQIAPRPPTPTGAPQOQSPAPAHVSHGP 180
DB      121  IAGSPAPAPKPTPMPPIAGSPAPPTPTESQIAPRPPTPTGAPQOQSPAPAHVSHGP 180
QY      121  IAGSPAPAPKPTPMPPIAGSPAPPTPTESQIAPRPPTPTGAPQOQSPAPAHVSHGP 180
DB      121  IAGSPAPAPKPTPMPPIAGSPAPPTPTESQIAPRPPTPTGAPQOQSPAPAHVSHGP 180
QY      181  HOPRRRTAPAPPAKMPIGSPAPSPASAPPTPMPPIAAGSPAPAPKPTPMPPIAGSPAP 240
DB      181  HOPRRRTAPAPPAKMPIGSPAPSPASAPPTPMPPIAAGSPAPAPKPTPMPPIAGSPAP 240
QY      241  GKVATGPSIOARLRABEASGAQLAPGTEPSAPLQGPSTIAPPTPAPTEPPSPSPOR 300
DB      241  GKVATGPSIOARLRABEASGAQLAPGTEPSAPLQGPSTIAPPTPAPTEPPSPSPOR 300
QY      301  NSGRARERVRHDLAQAHAAPDSITTAATGGRRRKRAAPDLIDATQKSLRPAKGPVK 360
DB      301  NSGRARERVRHDLAQAHAAPDSITTAATGGRRRKRAAPDLIDATQKSLRPAKGPVK 360
QY      361  KYKPKKATKPKPKVVSQSGMRHWALTRINIGLSPDEKYTELDLHARRRNRGSGYQA 420
DB      361  KYKPKKATKPKPKVVSQSGMRHWALTRINIGLSPDEKYTELDLHARRRNRGSGYQA 420
```


Db 361 KVKQKATKTPKPVVSQSGRMHWVHALTRINIGLSPEKYEELDHARVRNPRGSYQIA 420
QY 421 VVGLKGAGKTTLLTAALGSTLAQVRADRIILALDADPGAGNLADRVGOSGATTIADVIAEK 480
Db 421 VVGLKGAGKTTLLTAALGSTLAQVRADRIILALDADPGAGNLADRVGOSGATTIADVIAEK 480
QY 481 ELSHYNDIRAHTSVNAVNLEVLPAPEYSSAQRALSDADMHFIADPASFPYNLVIADCGAG 540
Db 481 ELSHYNDIRAHTSVNAVNLEVLPAPEYSSAQRALSDADMHFIADPASFPYNLVIADCGAG 540
QY 541 FPDPLTGVLSTVSGVVVAVSVSIDGAQOASVALDMLRNNGYODIARACVINHIMGE 600
Db 541 FPDPLTGVLSTVSGVVVAVSVSIDGAQOASVALDMLRNNGYODIARACVINHIMGE 600
QY 601 PNVAVKDLVHFEQOVQGRVVMVPMDRHIAAGTEISLDLDDPIYKKRVLELAALSDDF 660
Db 601 PNVAVKDLVHFEQOVQGRVVMVPMDRHIAAGTEISLDLDDPIYKKRVLELAALSDDF 660
QY 661 ERAGRR 666
Db 661 ERAGRR 666

RESULT 7
US-10-468-356-640
; Sequence 640, Application US/10468356
; Publication No. US20040197896A1
; GENERAL INFORMATION:
; APPLICANT: COLE, STEWART
; TITLE OF INVENTION: COMPARATIVE MYCOBACTERIAL GENOMICS AS A TOOL FOR
; TITLE OF INVENTION: IDENTIFYING TARGETS FOR THE DIAGNOSIS, PROPHYLAXIS OR
; FILE REFERENCE: 05394.0019
; CURRENT APPLICATION NUMBER: US/10/468,356
; CURRENT FILING DATE: 2003-08-19
; PRIOR APPLICATION NUMBER: 10/080,170
; PRIOR FILING DATE: 2002-02-22
; PRIOR APPLICATION NUMBER: 60/270,123
; PRIOR FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 655
; SOFTWARE: Patentln Ver. 3.2
; SEQ ID NO 640
; LENGTH: 666
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-10-468-356-640

Query Match 100.0%; Score 3552; DB 4; Length 666;
Best Local Similarity 100.0%; Pred. No. 2.5e-180;
Matches 666; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MAADYDKLFRPHGEMAPDDMAAOPFPDPSAPPASANLPEKNGQTPPTSDDLSEK 60
Db 1 MAADYDKLFRPHGEMAPDDMAAOPFPDPSAPPASANLPEKNGQTPPTSDDLSEK 60
QY 61 FVSAPPPEPPPPPTPTMPIAAGBPPSPBPASAKPTPTMPIAGBPPAPKPTPTMP 120
Db 61 FVSAPPPEPPPPPTPTMPIAAGBPPSPBPASAKPTPTMPIAGBPPAPKPTPTMP 120
QY 121 IAGBPPAPKPTPTMPIAAGBPPSPBPASAKPTPTMPIAGBPPAPKPTPTMP 180
Db 121 IAGBPPAPKPTPTMPIAAGBPPSPBPASAKPTPTMPIAGBPPAPKPTPTMP 180
QY 181 HOPRRATAPAPMAKMPGEBPPAPSRPSASBPAPPTPAOHSSRARGRKRTDTERNV 240
Db 181 HOPRRATAPAPMAKMPGEBPPAPSRPSASBPAPPTPAOHSSRARGRKRTDTERNV 240
QY 241 GKATGSGIOARLAEBSAQAOLAPGTBSPAPLQGRSYLAAPTRTAPTEPPSPSPOR 300
Db 241 GKATGSGIOARLAEBSAQAOLAPGTBSPAPLQGRSYLAAPTRTAPTEPPSPSPOR 300
QY 301 NSGRRAERRVHPDLAAQHAAPDSITAAITGGRRRRAAPDLATQKSLRPAKGPVK 360
Db 301 NSGRRAERRVHPDLAAQHAAPDSITAAITGGRRRRAAPDLATQKSLRPAKGPVK 360

Db 301 NSGRRAERRVHPDLAAQHAAPDSITAAITGGRRRRAAPDLATQKSLRPAKGPVK 360
QY 361 KVKQKATKTPKPVVSQSGRMHWVHALTRINIGLSPEKYEELDHARVRNPRGSYQIA 420
Db 361 KVKQKATKTPKPVVSQSGRMHWVHALTRINIGLSPEKYEELDHARVRNPRGSYQIA 420
QY 421 VVGLKGAGKTTLLTAALGSTLAQVRADRIILALDADPGAGNLADRVGOSGATTIADVIAEK 480
Db 421 VVGLKGAGKTTLLTAALGSTLAQVRADRIILALDADPGAGNLADRVGOSGATTIADVIAEK 480
QY 481 ELSHYNDIRAHTSVNAVNLEVLPAPEYSSAQRALSDADMHFIADPASFPYNLVIADCGAG 540
Db 481 ELSHYNDIRAHTSVNAVNLEVLPAPEYSSAQRALSDADMHFIADPASFPYNLVIADCGAG 540
QY 541 FPDPLTGVLSTVSGVVVAVSVSIDGAQOASVALDMLRNNGYODIARACVINHIMGE 600
Db 541 FPDPLTGVLSTVSGVVVAVSVSIDGAQOASVALDMLRNNGYODIARACVINHIMGE 600
QY 601 PNVAVKDLVHFEQOVQGRVVMVPMDRHIAAGTEISLDLDDPIYKKRVLELAALSDDF 660
Db 601 PNVAVKDLVHFEQOVQGRVVMVPMDRHIAAGTEISLDLDDPIYKKRVLELAALSDDF 660
QY 661 ERAGRR 666
Db 661 ERAGRR 666

RESULT 8
US-10-510-021-66
; Sequence 66, Application US/10510021
; Publication No. US20050220811A1
; GENERAL INFORMATION:
; APPLICANT: Cole, Stewart
; APPLICANT: Brosch, Roland
; APPLICANT: Brodin, Pricille
; APPLICANT: Majlessi, Laleh
; APPLICANT: Demangel, Caroline
; APPLICANT: Leclerc, Claude
; TITLE OF INVENTION: Identification of virulence associated regions RD1 and
; TITLE OF INVENTION: RD5 leading to improve vaccine of M. bovis BCG and M.
; FILE REFERENCE: D20217
; CURRENT APPLICATION NUMBER: US/10/510,021
; CURRENT FILING DATE: 2004-10-01
; PRIOR APPLICATION NUMBER: PCT/IB03/01789
; PRIOR FILING DATE: 2003-04-01
; PRIOR APPLICATION NUMBER: EP 02/290864
; PRIOR FILING DATE: 2002-04-05
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 66
; LENGTH: 666
; TYPE: PRT
; ORGANISM: mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: Protein sequence Rv3876
US-10-510-021-66

Query Match 100.0%; Score 3552; DB 5; Length 666;
Best Local Similarity 100.0%; Pred. No. 2.5e-180;
Matches 666; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MAADYDKLFRPHGEMAPDDMAAOPFPDPSAPPASANLPEKNGQTPPTSDDLSEK 60
Db 1 MAADYDKLFRPHGEMAPDDMAAOPFPDPSAPPASANLPEKNGQTPPTSDDLSEK 60
QY 61 FVSAPPPEPPPPPTPTMPIAAGBPPSPBPASAKPTPTMPIAGBPPAPKPTPTMP 120
Db 61 FVSAPPPEPPPPPTPTMPIAAGBPPSPBPASAKPTPTMPIAGBPPAPKPTPTMP 120
QY 121 IAGBPPAPKPTPTMPIAAGBPPSPBPASAKPTPTMPIAGBPPAPKPTPTMP 180
Db 121 IAGBPPAPKPTPTMPIAAGBPPSPBPASAKPTPTMPIAGBPPAPKPTPTMP 180

```
Db 121 IAGPEPAPKPPPPMPPIAGPAPTPTESQLAPRRPTPTQTTCGAPQPPSPAPHVSHGP 180
Qy 181 HQPRRTAPAPMAKMEIGEPSPAPSPSPAPSPTRPAPQHSRRARRGHRVTDTERNV 240
Db 181 HQPRRTAPAPMAKMEIGEPSPAPSPSPAPSPTRPAPQHSRRARRGHRVTDTERNV 240
Qy 241 GKVATGPGSIQARIABEASGAQLAPGTSPAPVLCGPSTLAPPTPAPTEPPSPSPOR 300
Db 241 GKVATGPGSIQARIABEASGAQLAPGTSPAPVLCGPSTLAPPTPAPTEPPSPSPOR 300
Qy 301 NSGRARRRVHPDLAQHAAOPDSITTAATTGRRRRRAAPDLATQKSLRPAKGPVKY 360
Db 301 NSGRARRRVHPDLAQHAAOPDSITTAATTGRRRRRAAPDLATQKSLRPAKGPVKY 360
Qy 361 KYKPKPKATKPKPKVVSQSGRMHVHATRTINLGLSPDEKYEELDLHARVRNPGSYQIA 420
Db 361 KYKPKPKATKPKPKVVSQSGRMHVHATRTINLGLSPDEKYEELDLHARVRNPGSYQIA 420
Qy 421 VVGLKGAAGKTTTLTAALGSTLAQVRADRIIALDADPGAGNLADRVGOSGATTADYLAER 480
Db 421 VVGLKGAAGKTTTLTAALGSTLAQVRADRIIALDADPGAGNLADRVGOSGATTADYLAER 480
Qy 481 ELSHYNDIRAHTSVNAVNLEVLPAPEYSSAQORALSDADMHFIADPASRFTNLVLADCGAG 540
Db 481 ELSHYNDIRAHTSVNAVNLEVLPAPEYSSAQORALSDADMHFIADPASRFTNLVLADCGAG 540
Qy 541 FPDPLTRGVLTSTVSGVVVVASVSDGAQOASVALDMLRNNGYODLASRAACVNIHIMP 600
Db 541 FPDPLTRGVLTSTVSGVVVVASVSDGAQOASVALDMLRNNGYODLASRAACVNIHIMP 600
Qy 601 PNVAVKDLVHFEQOVQPGRVVVMPPMDRHIAGTEISLDLDDPIYKRVLELAALSD 660
Db 601 PNVAVKDLVHFEQOVQPGRVVVMPPMDRHIAGTEISLDLDDPIYKRVLELAALSD 660
Qy 661 ERAGRR 666
Db 661 ERAGRR 666

RESULT 9
US-09-791-171-94
; Sequence 94, Application US/09791171
; Patent No. US200209436A1
; GENERAL INFORMATION:
; APPLICANT: ANDERSEN, Peter
; APPLICANT: NIELSEN, Rikke
; APPLICANT: OETTINGER, Thomas
; APPLICANT: RASMUSSEN, Peter Birk
; APPLICANT: ROSENKRANDS, Ida
; APPLICANT: WELDINGH, Karin
; APPLICANT: FLORIO, Walter
; TITLE OF INVENTION: NUCLEIC ACIDS FRAGMENTS AND POLYPEPTIDE FRAGMENTS
; TITLE OF INVENTION: DERIVED FROM M. TUBERCULOSIS
; FILE REFERENCE: 670001-2002.1
; CURRENT APPLICATION NUMBER: US/09/791,171
; CURRENT FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 09/050,739
; PRIOR FILING DATE: 1998-03-30
; PRIOR APPLICATION NUMBER: 0376/97
; PRIOR FILING DATE: 1997-04-02
; PRIOR APPLICATION NUMBER: 1277/97
; PRIOR FILING DATE: 1997-11-10
; PRIOR APPLICATION NUMBER: 60/044,624
; PRIOR FILING DATE: 1997-04-18
; PRIOR APPLICATION NUMBER: 60/070,488
; PRIOR FILING DATE: 1998-01-05
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 94
; LENGTH: 308
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-09-791-171-94
```

```
Query Match 44.3%; Score 1575; DB 3; Length 308;
Best Local Similarity 99.7%; Pred. No. 6,9e-76;
Matches 307; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 359 VKKVPQKPKATKPKPKVVSQSGRMHVHATRTINLGLSPDEKYEELDLHARVRNPGSYQ 418
Db 1 MKKVPQKPKATKPKPKVVSQSGRMHVHATRTINLGLSPDEKYEELDLHARVRNPGSYQ 60
Qy 419 IAVVGLKGAAGKTTTLTAALGSTLAQVRADRIIALDADPGAGNLADRVGOSGATTADYLA 478
Db 61 IAVVGLKGAAGKTTTLTAALGSTLAQVRADRIIALDADPGAGNLADRVGOSGATTADYLA 120
Qy 479 ELSHYNDIRAHTSVNAVNLEVLPAPEYSSAQORALSDADMHFIADPASRFTNLVLADCG 538
Db 121 ELSHYNDIRAHTSVNAVNLEVLPAPEYSSAQORALSDADMHFIADPASRFTNLVLADCG 180
Qy 539 AGFPDLTRGVLTSTVSGVVVVASVSDGAQOASVALDMLRNNGYODLASRAACVNIHIMP 598
Db 181 AGFPDLTRGVLTSTVSGVVVVASVSDGAQOASVALDMLRNNGYODLASRAACVNIHIMP 240
Qy 599 GEPNVAVKDLVHFEQOVQPGRVVVMPPMDRHIAGTEISLDLDDPIYKRVLELAALSD 658
Db 241 GEPNVAVKDLVHFEQOVQPGRVVVMPPMDRHIAGTEISLDLDDPIYKRVLELAALSD 300

RESULT 10
US-09-804-980-94
; Sequence 94, Application US/09804980
; Publication No. US20030147897A1
; GENERAL INFORMATION:
; APPLICANT: Statens Serum Institut
; APPLICANT: Anderson, Peter
; TITLE OF INVENTION: M. tuberculosis Antigens
; FILE REFERENCE: 670001-2002.4
; CURRENT APPLICATION NUMBER: US/09/804,980
; CURRENT FILING DATE: 2001-03-12
; NUMBER OF SEQ ID NOS: 257
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 94
; LENGTH: 308
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-09-804-980-94

Query Match 44.3%; Score 1575; DB 3; Length 308;
Best Local Similarity 99.7%; Pred. No. 6,9e-76;
Matches 307; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 359 VKKVPQKPKATKPKPKVVSQSGRMHVHATRTINLGLSPDEKYEELDLHARVRNPGSYQ 418
Db 1 MKKVPQKPKATKPKPKVVSQSGRMHVHATRTINLGLSPDEKYEELDLHARVRNPGSYQ 60
Qy 419 IAVVGLKGAAGKTTTLTAALGSTLAQVRADRIIALDADPGAGNLADRVGOSGATTADYLA 478
Db 61 IAVVGLKGAAGKTTTLTAALGSTLAQVRADRIIALDADPGAGNLADRVGOSGATTADYLA 120
Qy 479 ELSHYNDIRAHTSVNAVNLEVLPAPEYSSAQORALSDADMHFIADPASRFTNLVLADCG 538
Db 121 ELSHYNDIRAHTSVNAVNLEVLPAPEYSSAQORALSDADMHFIADPASRFTNLVLADCG 180
Qy 539 AGFPDLTRGVLTSTVSGVVVVASVSDGAQOASVALDMLRNNGYODLASRAACVNIHIMP 598
Db 181 AGFPDLTRGVLTSTVSGVVVVASVSDGAQOASVALDMLRNNGYODLASRAACVNIHIMP 240
Qy 599 GEPNVAVKDLVHFEQOVQPGRVVVMPPMDRHIAGTEISLDLDDPIYKRVLELAALSD 658
Db 241 GEPNVAVKDLVHFEQOVQPGRVVVMPPMDRHIAGTEISLDLDDPIYKRVLELAALSD 300
```

QY 659 DPERAGR 666
DB 301 DPERAGR 308

RESULT 11

US-10-620-246-94
; Sequence 94, Application US/10620246
; Publication No. US20040115211A1
; GENERAL INFORMATION:
; APPLICANT: ANDERSEN, Peter
; APPLICANT: NIELSEN, Rikke
; APPLICANT: OETTINGER, Thomas
; APPLICANT: RASMUSSEN, Peter Blak
; APPLICANT: ROSENKRANDS, Ida
; APPLICANT: WELDINGH, Karin
; APPLICANT: FLORIO, Walter
; TITLE OF INVENTION: NUCLEIC ACIDS FRAGMENTS AND POLYPEPTIDE FRAGMENTS
; TITLE OF INVENTION: DERIVED FROM M. TUBERCULOSIS
; FILE REFERENCE: 670001-2002.1A
; CURRENT FILING DATE: 2003-07-15
; PRIOR APPLICATION NUMBER: 09/050,739
; PRIOR FILING DATE: 1998-03-30
; PRIOR APPLICATION NUMBER: 0376/97
; PRIOR FILING DATE: 1997-04-02
; PRIOR APPLICATION NUMBER: 1277/97
; PRIOR FILING DATE: 1997-11-10
; PRIOR APPLICATION NUMBER: 60/044,624
; PRIOR FILING DATE: 1997-04-18
; PRIOR APPLICATION NUMBER: 60/070,488
; PRIOR FILING DATE: 1998-01-05
; PRIOR APPLICATION NUMBER: 10/138,473
; PRIOR FILING DATE: 2002-05-02
; PRIOR APPLICATION NUMBER: 09/791,171
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 09/415,884
; PRIOR FILING DATE: 1999-10-08
; PRIOR APPLICATION NUMBER: 60/116,673
; PRIOR FILING DATE: 1999-01-21
; PRIOR APPLICATION NUMBER: 1281/98
; PRIOR FILING DATE: 1998-10-08
; NUMBER OF SEQ ID NOS: 173
; SOFTWARE: PatentIn Ver. 2.0
; LENGTH: 308
; SEQ ID NO 94
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-10-620-246-94

Query Match 44.3%; Score 1575; DB 4; Length 308;
Best Local Similarity 99.7%; Pred. No. 6.9e-76;
Matches 307; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 359 VKKVKQKPKATPKPVVSGRGWRHWHALTRINIGLSPDEKYLELHARVRNPRGSYQ 418
DB 1 MKKVKQKPKATPKPVVSGRGWRHWHALTRINIGLSPDEKYLELHARVRNPRGSYQ 60
QY 419 IAVVKGKAGAKTTLTALAGSTLAQVRADRIILALDADPGAGNLADYRGOSGATIDVLA 478
DB 61 IAVVKGKAGAKTTLTALAGSTLAQVRADRIILALDADPGAGNLADYRGOSGATIDVLA 120
QY 479 EKELSHYNDIRATSVNAVNLAVLPAPVSSAQRALSDADMHFTADPASFFYLVLADCG 538
DB 121 EKELSHYNDIRATSVNAVNLAVLPAPVSSAQRALSDADMHFTADPASFFYLVLADCG 180
QY 539 AGFPDLRTGVLSTVSGVVVAVSVSDGAQASVALDMLNNGYODLASPAVCVNIHNP 598
DB 181 AGFPDLRTGVLSTVSGVVVAVSVSDGAQASVALDMLNNGYODLASPAVCVNIHNP 240
QY 599 GERNVAVKDLVNHFEQOVQGRVYVVPMDRHIAAGTBSLDLDPYIKKVTLELAALSD 658
DB 241 GERNVAVKDLVNHFEQOVQGRVYVVPMDRHIAAGTBSLDLDPYIKKVTLELAALSD 300

QY 659 DPERAGR 666
DB 301 DPERAGR 308

RESULT 12

US-10-080-170-10
; Sequence 10, Application US/10080170
; Publication No. US20030129601A1
; GENERAL INFORMATION:
; APPLICANT: COLE, S.T.
; TITLE OF INVENTION: COMPARATIVE MYCOBACTERIAL GENOMICS AS A TOOL FOR
; TITLE OF INVENTION: IDENTIFYING TARGETS FOR THE DIAGNOSIS, PROPHYLAXIS OR
; FILE REFERENCE: 03495.0218
; CURRENT APPLICATION NUMBER: US/10/080,170
; PRIOR FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: 60/270,123
; PRIOR FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 652
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 586
; TYPE: PRT
; ORGANISM: Mycobacterium leprae
US-10-080-170-10

Query Match 43.6%; Score 1550; DB 4; Length 586;
Best Local Similarity 49.2%; Pred. No. 2.9e-74;
Matches 335; Conservative 66; Mismatches 168; Indels 112; Gaps 15;

QY 1 MAADYDKLFRPHGEMAPDDMAAQPFPDSAPPAASANTLKPNGQTPPTSDLSER 60
DB 1 MAADYDKLFRPDGGAASPDQAAGLFDAPLPPPIPTCTTPNGVSP-MPWSQ 59
QY 61 FVSAPPPPPPPPTPMPPIAGBPSPBPASKPPTPMPDIAGBPAPKPPPTPMP 120
DB 60 L-----PPNPASKSLPMPPI--GSSVQPPPASSPAPMVSAP--PRSPASLMP 110
QY 121 IAGEPAP-----KPTTPMPPIAGPAPTESGLAPRRPTPTGAPQPE 169
DB 111 ISEPPQMPAPAEHQFAKAPPSVPIPIPEBP-----ACAPATPMTPTDGSOR-- 161
QY 170 SPAPHVSHGHPQRTAPAPPAKMPIGPP--PAPS--RPSAPABEPTPAPQHSR 224
DB 162 TPV-----TSPEPSLAPPE-AOPATPSPILPRMSPBPAPRSANQHSR 207
QY 225 RARGRHYRTDERNVKVAATGPSIQARLAREASGAOLAPGTBPAPAPGQPSYLAAP 284
DB 208 HARGGHHHDETOQANPASATBPMIAPAKTAE--LRQPHAAABPAP----- 253
QY 285 TRAPTEPPPSBPQNSGRBARRRVHPDLAQAHAADPSTIAATTGGRRRGAAPDLD 344
DB 254 -----TQHLTRPDGLVSHRTALDLS-TATSAIGVQGRST-- 287
QY 345 ATOKSLRAPAKGPKVKVQKPKATPKPVVSGRGWRHWHALTRINIGLSPDEKYLEL 404
DB 288 -----GAKKPSVAAKRRHWHALTRINIGLSPDEKYLEL 384
QY 405 LHAVRNPRGSYOIYAVGKAGAKTTLTALAGSTLAQVRADRIILALDADPGAGNLADR 464
DB 325 LRRVRRPGRGSQIILGKAGAKTTLTALAGSTLAQVRADRIILVVDADTSCGNLADR 384
QY 465 VERQSGATTADYAEKELSHYNDIRATSVNAVNLAVLPAPVSSAQRALSDADMHFTAD 524
DB 385 AGFSPSANTADLADADVSVSYNDIRHTSVNANLEVLPAEYSTAQHLSGSDMNPFAA 444
QY 525 PASRFNVLADCGAGFPDLRTGVLSTVSGVVVAVSVSDGAQASVALDMLNNGYOD 584
DB 445 TVSKTYNVLADCGVGLFDPVTRGVLSTASGVVIVSTVDAAKQALADMLNNGYOD 504
QY 585 LASRACCVINHIMPGRNVAVDLVNHFEQOVQGRVYVVPMDRHIAAGTBSLDLDPY 644

Db 505 LLSRACVIVNHVWPKBPNIASKOLVQOFEQOIQPGHVVLVPMWKHIAAGTEIRLIDLPL 564
QY 645 YKRKVLTELAAALSDPERAGR 665
Db 565 YRRRIELAAALSDPERAGR 585

RESULT 13

US-10-282-122A-63640
Sequence 63640, Application US/10282122A
Publication No. US20040029129A1
GENERAL INFORMATION:
APPLICANT: Wang, Liangu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Habelbeck, Robert
APPLICANT: Ohlsen, Karl
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Foreyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELIRA.034A
CURRENT APPLICATION NUMBER: US/10/282.122A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See file Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: PatentIn version 3.1
SEQ ID NO 63640
LENGTH: 586
TYPE: PRT
ORGANISM: Mycobacterium leprae
US-10-282-122A-63640

Query Match 43.6%; Score 1550; DB 4; Length 586;

Best Local Similarity 49.2%; Pred. No. 2.9e-74;
Matches 335; Conservative 66; Mismatches 168; Indels 112; Gaps 15;

QY 1 MAADVYKLRFRPHGEMAPDDMAQPFDPSPAPASANLPKPNQGPPTPPSDLSER 60
Db 1 MAADVYKLRFLDGAIVASPDQAABQLFDDAFLYPPPIPTCTTPNGEIVASP-MPDWSQ 59
QY 61 FVSAPPPEPPPPPTPMPPIAAGBPSPBPASAKPTPPMPIAGBEPAPPKPTPPMP 120
Db 60 L-----PFPNPASAKSLPMPPI--GSSVQPPPASSSSPAPMPVSA---PSPASASLP 110
QY 121 IAGBEPAP-----KPTPPMPIAGPAFTPTESQALAPRPPTPTPTGAPQOPE 169
Db 111 ISBPPQWPPPAEAPHQFAKABPPSVPIPIBPSP-----AKPATPMPTPTDGSOR-- 161

QY 170 SPAPHVSHGPQPKRTAPAPWAKPIGBPP--PAPS---PSPASAPPTPRAPQHSR 224
Db 162 TVV-----TSPEPSIAEPE-AQPPATPKPSLLPBPMSPPAPBPSPANQHSR 207
QY 225 PARRGHRTYDTERVAVGKATGTPSTQALRAEASGAQALAPGTSPAPALGQPRSLAP 284
Db 208 HARGHHHHDFTQOANPASATEPMTAPARTAE---LRQAPHAABPAP----- 253
QY 285 TRPAPTEPPSPSPQNSGRABRRVHPDLAQAHAACPDSITTAATTGRRRRKRAPDLD 344
Db 254 -----TQHLTRPDGLVSHRTALHDS-TAISAIGVQGRST----- 287
QY 345 ATQKSLRPAKGPVKVQKVPQKPKATKPKVVSQSGWRHVAALTRINIGLSPDEXEYLD 404
Db 288 -----GAKKPSKVAAKGMHVMHTVTRINIGLSPDEREYLD 324
QY 405 LPARVRNPRGSYQIAYVGLKGGAGKTTITALLGSTLAQVNRDRIALDADFGAGTLAR 464
Db 325 LRTVRVRPPRGSGYQIGLIGLKGAGKTVTVLGSMPARVRNDRILVVDAPSCGLADR 384
QY 465 VGRSGATTADVLAEBKLSHYNDIRAHTSVNAVNLVLPAPEYSSAQRALSDADMHFIAD 524
Db 385 AGPSEANITADLDADQVASTNDIRHITSVNAVNLVLPAPEYSTRQHALSGEDNFPAA 444
QY 525 PASRFYNTVLADCGAGFPDPLTRGYLSTVSGVVVAVSIDAQAQASVALDWLRNNGYOD 584
Db 445 TVSKYTNVWLADQVGLFDPVTRGVLTASGVVITVSTSVDAARQAALDWLRHNGYOD 504
QY 585 LLSRACVIVNHVWPKBPNIASKOLVQOFEQOIQPGHVVLVPMWKHIAAGTEIRLIDLPL 644
Db 505 LLSRACVIVNHVWPKBPNIASKOLVQOFEQOIQPGHVVLVPMWKHIAAGTEIRLIDLPL 564
QY 645 YKRKVLTELAAALSDPERAGR 665
Db 565 YRRRIELAAALSDPERAGR 585

RESULT 14

US-10-080-170-10
Sequence 10, Application US/10080170
Publication No. US20040121322A9
GENERAL INFORMATION:
APPLICANT: COLE, S.T.
TITLE OF INVENTION: COMPARATIVE MYCOBACTERIAL GENOMICS AS A TOOL FOR
IDENTIFYING TARGETS FOR THE DIAGNOSIS, PROPHYLAXIS OR
TREATMENT OF MYCOBACTERIOSES
FILE REFERENCE: 03495.0218
CURRENT APPLICATION NUMBER: US/10/080.170
CURRENT FILING DATE: 2002-06-10
PRIOR APPLICATION NUMBER: 60/270,123
PRIOR FILING DATE: 2001-02-22
NUMBER OF SEQ ID NOS: 652
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 10
LENGTH: 586
TYPE: PRT
ORGANISM: Mycobacterium leprae
US-10-080-170-10

Query Match 43.6%; Score 1550; DB 4; Length 586;

Best Local Similarity 49.2%; Pred. No. 2.9e-74;
Matches 335; Conservative 66; Mismatches 168; Indels 112; Gaps 15;

QY 1 MAADVYKLRFRPHGEMAPDDMAQPFDPSPAPASANLPKPNQGPPTPPSDLSER 60
Db 1 MAADVYKLRFLDGAIVASPDQAABQLFDDAFLYPPPIPTCTTPNGEIVASP-MPDWSQ 59
QY 61 FVSAPPPEPPPPPTPMPPIAAGBPSPBPASAKPTPPMPIAGBEPAPPKPTPPMP 120
Db 60 L-----PFPNPASAKSLPMPPI--GSSVQPPPASSSSPAPMPVSA---PSPASASLP 110
QY 121 IAGBEPAP-----KPTPPMPIAGPAFTPTESQALAPRPPTPTPTGAPQOPE 169

```

Db      111  ISEPPQWPAABEHPQAKAEPSPVPPIPIHEPSP-----AKPATMPTPIIDGSR-- 161
Qy      170  SPAHVHSHGHQRRRTAPAPPMAMKPIGEP--PAPS---RPSASAPPTBPAPQHSR 224
Db      162  TPV-----TSPEPSLAEFE-AQPATPKKSLPRPMSSPPEAPRPSANQHSR 207
Qy      225  RARGHRRTDTERNVKVAATGPSIOARLAEBSGAQLAPGTEPSPAPLQPPSYLAPP 284
Db      208  HAARGHHHROBTQANPASATEBPWIAPRARTAE---LRQAPHAAAEAP----- 253
Qy      285  TRAPATPPSPSPQNRSGRAERVRVPDLAAGHAAQPPSITPAATTGRRRKAAPDLD 344
Db      254  -----TQHLTRPDGLVSHRTLHDS-TATSAIGVQTGST----- 287
Qy      345  ATOKSLRPAKGPVKKKVKQKPKATYPPKVVSGQRGRHWVHALTRINLGLSPDEKYELD 404
Db      288  -----GAKKPSKVVAKRGRHWVHTVTRINLGLSPDERYELD 324
Qy      405  LHAARBNPRGSYQIAVVGKKGAGKTTTLTAALGSTIAOVRADRIILALDADPGAGNLADR 464
Db      325  LRTRVRPRPGSYQIGILGKKGAGKTTVTVLGSMPARVRNDRILVVDADTSCGNLADR 384
Qy      465  VGRSGATTIDVLAEKELSHYNDIRATTSVNAVNLBYLPAPEYSSAQALSDADWHFIAD 524
Db      385  AGRFSEANIIDLADKDVKSYNDRITRTSVNAVNLBYLPAEYSTAQHALLSGEDMNFALAA 444
Qy      525  PASRFYNLVLADGAGFPDPLTRGVLTSGVVVVAVSIDGAQAQSVALLDMLNNQYOD 584
Db      445  TVSKTYVNLADCGVGLFDPVTRGVLTSGVIVTVTSVDAARQAAILDMLRHNGYOD 504
Qy      585  LASRACVIVNHIMGEENNAVKDLVRHFEQOVQPGRVVMPWDRHIAAGTEISIDLDP1 644
Db      505  LLSRACVIVNHIMWPKENINASKDLVQGFEOQIQGRVVVLEPMDGHIAAGTEIRIDRLDPL 564
Qy      645  YKRKVTLELAALSDDFERRAGR 665
Db      565  YRRRIELELAALSDDFERRAGR 585

RESULT 15
US-10-468-356-10
; Sequence 10, Application US/10468356
; Publication No. US20040197896A1
; GENERAL INFORMATION:
; APPLICANT: COLE, STEWART
; TITLE OF INVENTION: COMPARATIVE MYCOBACTERIAL GENOMICS AS A TOOL FOR
; TITLE OF INVENTION: IDENTIFYING TARGETS FOR THE DIAGNOSIS, PROPHYLAXIS OR
; FILE REFERENCE: 05394.0019
; CURRENT APPLICATION NUMBER: US/10/468,356
; PRIOR FILING DATE: 2003-08-19
; PRIOR APPLICATION NUMBER: 10/080,170
; PRIOR FILING DATE: 2002-02-22
; PRIOR APPLICATION NUMBER: 60/270,123
; NUMBER OF SEQ ID NOS: 655
; SOFTWARE: Patent Ver. 3.2
; SEQ ID NO 10
; LENGTH: 586
; TYPE: PRT
; ORGANISM: Mycobacterium leprae
US-10-468-356-10

Query Match      43.6%; Score 1550; DB 4; Length 586;
Best Local Similarity 49.2%; Pred. No. 2.9e-74;
Matches 335; Conservative 66; Mismatches 168; Indels 112; Gaps 15;
```

```

Qy      1  MAADYDCLFRPHGMEAPDDMAAPFPDPSPASFPAPASANLPRKNGOTPPPTSDLSER 60
Db      1  MAADYDCLFRIDDGAVASPDQAAEQLPDADLVPPIIPTCTTTPNGEVAAP-MPDWSQ 59
Qy      61  FVSAFPPPPPPPPPTPMPIAAGEPPSPASKEPTTPMPIAGPEPAPPPKPTTPMP 120
```

```

Db      60  L---PENPPAASKSPLEPMTI--GSSVQPPPPASSSPRAMVVSAP---PRSPASLMP 110
Qy      121  IAGEPEAP-----KPTPPMPIAGPATPTPESOLAPPRPPTPOTPTGAPQOPE 169
Db      111  ISEPPQWPAABEHPQAKAEPSPVPPIPIHEPSP-----AKPATMPTPIIDGSR-- 161
Qy      170  SPAHVHSHGHQRRRTAPAPPMAMKPIGEP--PAPS---RPSASAPPTBPAPQHSR 224
Db      162  TPV-----TSPEPSLAEFE-AQPATPKKSLPRPMSSPPEAPRPSANQHSR 207
Qy      225  RARGHRRTDTERNVKVAATGPSIOARLAEBSGAQLAPGTEPSPAPLQPPSYLAPP 284
Db      208  HAARGHHHROBTQANPASATEBPWIAPRARTAE---LRQAPHAAAEAP----- 253
Qy      285  TRAPATPPSPSPQNRSGRAERVRVPDLAAGHAAQPPSITPAATTGRRRKAAPDLD 344
Db      254  -----TQHLTRPDGLVSHRTLHDS-TATSAIGVQTGST----- 287
Qy      345  ATOKSLRPAKGPVKKKVKQKPKATYPPKVVSGQRGRHWVHALTRINLGLSPDEKYELD 404
Db      288  -----GAKKPSKVVAKRGRHWVHTVTRINLGLSPDERYELD 324
Qy      405  LHAARBNPRGSYQIAVVGKKGAGKTTTLTAALGSTIAOVRADRIILALDADPGAGNLADR 464
Db      325  LRTRVRPRPGSYQIGILGKKGAGKTTVTVLGSMPARVRNDRILVVDADTSCGNLADR 384
Qy      465  VGRSGATTIDVLAEKELSHYNDIRATTSVNAVNLBYLPAPEYSSAQALSDADWHFIAD 524
Db      385  AGRFSEANIIDLADKDVKSYNDRITRTSVNAVNLBYLPAEYSTAQHALLSGEDMNFALAA 444
Qy      525  PASRFYNLVLADGAGFPDPLTRGVLTSGVVVVAVSIDGAQAQSVALLDMLNNQYOD 584
Db      445  TVSKTYVNLADCGVGLFDPVTRGVLTSGVIVTVTSVDAARQAAILDMLRHNGYOD 504
Qy      585  LASRACVIVNHIMGEENNAVKDLVRHFEQOVQPGRVVMPWDRHIAAGTEISIDLDP1 644
Db      505  LLSRACVIVNHIMWPKENINASKDLVQGFEOQIQGRVVVLEPMDGHIAAGTEIRIDRLDPL 564
Qy      645  YKRKVTLELAALSDDFERRAGR 665
Db      565  YRRRIELELAALSDDFERRAGR 585
```

Search completed: April 14, 2006, 18:40:36
Job time : 238.816 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using SW model

Run on: April 14, 2006, 18:37:38 ; Search time 28.5366 Seconds
(without alignments)
991.565 Million cell updates/sec

Title: US-10-620-246-70

Perfect score: 3552
Sequence: 1 MADYDYLFRPHGMEAPD.....KYLALALSDPERAGR 666

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 217505 seqs, 42489236 residues

Total number of hits satisfying chosen parameters: 217505

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA New:
1: /SIDS5/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
2: /SIDS5/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
3: /SIDS5/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
4: /SIDS5/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
5: /SIDS5/ptodata/2/pubpaa/US05_NEW_PUB.pep.*
6: /SIDS5/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
7: /SIDS5/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
8: /SIDS5/ptodata/2/pubpaa/US60_NEW_PUB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed.
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	444	12.5	1000	US-11-036-256-95	Sequence 95, Appl
2	444	12.5	1001	US-11-036-256-5	Sequence 5, Appl
3	444	12.5	1001	US-11-036-256-96	Sequence 96, Appl
4	388	10.9	717	US-11-096-568A-28799	Sequence 28799, A
5	373.5	10.6	694	US-11-096-568A-29818	Sequence 29818, A
6	366.5	10.3	411	US-11-036-256-89	Sequence 89, Appl
7	359.5	10.1	681	US-11-096-568A-2501	Sequence 2501, A
8	348	9.8	346	US-11-036-256-33	Sequence 33, Appl
9	347	9.8	941	US-11-169-232-14	Sequence 14, Appl
10	347	9.8	1022	US-11-169-232-84	Sequence 84, Appl
11	347	9.8	1038	US-11-169-232-74	Sequence 74, Appl
12	347	9.8	1049	US-11-169-232-58	Sequence 58, Appl
13	347	9.8	1140	US-11-169-232-104	Sequence 104, Appl
14	347	9.8	1270	US-11-169-232-44	Sequence 44, Appl
15	347	9.8	1311	US-11-169-232-42	Sequence 42, Appl
16	347	9.8	1313	US-11-169-232-142	Sequence 142, Appl
17	347	9.8	1314	US-11-169-232-50	Sequence 50, Appl
18	347	9.8	1330	US-11-169-232-46	Sequence 46, Appl
19	347	9.8	1354	US-11-169-232-60	Sequence 60, Appl
20	347	9.8	1361	US-11-169-232-48	Sequence 48, Appl
21	347	9.8	1361	US-11-169-232-40	Sequence 40, Appl
22	347	9.8	1363	US-11-169-232-52	Sequence 52, Appl
23	347	9.8	1404	US-11-169-232-52	Sequence 2, Appl
24	347	9.8	1404	US-11-169-232-62	Sequence 62, Appl
25	341	9.6	1267	US-11-096-568A-28960	Sequence 28960, A

26	341	9.6	1274	US-11-096-568A-28959	Sequence 28959, A
27	341	9.6	1480	US-11-096-568A-28958	Sequence 28958, A
28	337	9.5	895	US-11-096-568A-29626	Sequence 29626, A
29	328	9.2	639	US-11-096-568A-27757	Sequence 27757, A
30	328	9.2	674	US-11-096-568A-27756	Sequence 27756, A
31	328	9.2	786	US-11-096-568A-27755	Sequence 27755, A
32	323	9.1	797	US-10-330-773-663	Sequence 663, App
33	322.5	9.1	647	US-11-087-099-7887	Sequence 7887, App
34	322.5	9.1	1264	US-11-188-298-18316	Sequence 18316, A
35	321	9.0	1264	US-11-096-568A-30287	Sequence 30287, A
36	321	9.0	1301	US-11-096-568A-30286	Sequence 30286, A
37	321	9.0	1307	US-11-096-568A-30285	Sequence 30285, A
38	317.5	8.9	406	US-11-096-568A-12552	Sequence 12552, A
39	315.5	8.9	1560	US-10-330-773-661	Sequence 661, App
40	313.5	8.8	395	US-11-108-172-1068	Sequence 1068, App
41	306.5	8.6	5179	US-11-036-256-31	Sequence 31, Appl
42	304.5	8.6	228	US-11-036-256-158	Sequence 158, App
43	304	8.6	101	US-11-036-256-159	Sequence 159, App
44	298	8.4	1823	US-10-995-561-988	Sequence 988, App
45					

ALIGNMENTS

```
RESULT 1
US-11-036-256-95
; Sequence 95, Application US/11036256
; Publication NO. US20060026719A1
; GENERAL INFORMATION:
; APPLICANT: KIELISZEWSKI, MARCIA
; APPLICANT: XU, JIANFENG
; TITLE OF INVENTION: METHODS OF PRODUCING PEPTIDES/PROTEINS IN PLANTS AND
; TITLE OR INVENTION: PEPTIDES/PROTEINS PRODUCED THEREBY
; FILE REFERENCE: 27211/04130
; CURRENT APPLICATION NUMBER: US/11/036,256
; PRIOR FILING DATE: 2005-01-14
; PRIOR APPLICATION NUMBER: 60/602,562
; PRIOR FILING DATE: 2004-08-18
; PRIOR APPLICATION NUMBER: 60/582,027
; PRIOR FILING DATE: 2004-06-22
; PRIOR APPLICATION NUMBER: 60/536,486
; PRIOR FILING DATE: 2004-01-14
; NUMBER OF SEQ ID NOS: 173
; SOFTWARE: PatentIn Ver. 3.3
; SEQ ID NO 95
; LENGTH: 1000
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURES:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: amino acid motif
; OTHER INFORMATION: This sequence may encompass 2-1,000 residues of Pro
US-11-036-256-95

Query Match      12.5%; Score 444; DB 7; Length 1000;
Best Local Similarity 35.5%; Pred. No. 4.4e-18;
Matches 94; Conservative 0; Mismatches 115; Indels 56; Gaps 2;

Cy 34 PPAPAAANTLPEKPGQTPPTSDLSRFVSAPPPPPPPPTPMPPIAGBPPSPDEPA 93
Db 1 PPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPP 48

Cy 94 ASKPPPPMPPIAGBPPAPKPPPTPMPPIAGBPPAPKPPPTPMPPIAGBPPPTBSQLAPP 153
Db 49 PPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPP 108

Cy 154 RPPPTQPTGCAQQQSPAPAHVPSHGPHQRRKAPAPAPMAKMGIGSPAPSPSPASPAE 213
Db 109 PPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPP 168

Cy 214 PPTRPAPVQHSRRARRRGHRYRTDTERNVGKATGPSICARLAEASGAQLAPGTSPSPAP 273
```

Db 169 PPPPPPP-----PPPPPPPP 184
QY 274 LQGRSYLAPTRPAPTEPPSPSP 298
Db 185 PPPPPPPPPPPPPPPPPPPPPPP 209

RESULT 2

US-11-036-256-5
/ Sequence 5, Application US/11036256
/ Publication No. US20060026719A1
/ GENERAL INFORMATION:
/ APPLICANT: KIELISZEWSKI, MARCIA
/ APPLICANT: XU, JIANPENG
/ TITLE OF INVENTION: METHODS OF PRODUCING PEPTIDES/PROTEINS IN PLANTS AND
/ FILE REFERENCES: 27211/04130
/ CURRENT APPLICATION NUMBER: US/11/036,256
/ PRIOR FILING DATE: 2005-01-14
/ PRIOR APPLICATION NUMBER: 60/602,562
/ PRIOR FILING DATE: 2004-08-18
/ PRIOR APPLICATION NUMBER: 60/582,027
/ PRIOR FILING DATE: 2004-06-22
/ PRIOR APPLICATION NUMBER: 60/536,486
/ PRIOR FILING DATE: 2004-01-14
/ NUMBER OF SEQ ID NOS: 173
/ SOFTWARE: PatentIn Ver. 3.3
/ SEQ ID NO 5
/ LENGTH: 1001
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence: Synthetic
/ NAME/KEY: MOD_RES
/ LOCATION: (1)
/ OTHER INFORMATION: Lys, Ser, Ala, Thr, Gly, or Val
/ NAME/KEY: MOD_RES
/ LOCATION: (2)..(1001)
/ OTHER INFORMATION: This region may encompass 2 to 1,000 residues of
/ OTHER INFORMATION: Proline
US-11-036-256-5

Query Match 12.5%; Score 444; DB 7; Length 1001;
Best Local Similarity 35.5%; Pred. No. 4,4e-18;
Matches 94; Conservative 0; Mismatches 115; Indels 56; Gaps 2;
QY 34 PPAPASANI.PKPNQTPPTSDLSERFVSAPPPPPPPPPPTPMPPIAGSPSPSPA 93
Db 2 PPP 49
QY 94 ASKPPTPMPPIAGPBEAPKPPPTPMPPIAGPBEAPKPPPTPMPPIAGPAPPTESQLAPP 153
Db 50 PPP 109
QY 154 RPPPTGTPGAGQQPSSPAHVSHGPHQRRTAPAPMAKMPIGSPSPSPSPSPA 213
Db 110 PPP 169
QY 214 PPTRPAPQHSRRARRGHRYTDTERNVKGATGTSIQARLRABEASGAQLAPGTBSPAP 273
Db 170 PPPPPPP-----PPPPPPPP 185
QY 274 LQGRSYLAPTRPAPTEPPSPSP 298
Db 186 PPPPPPPPPPPPPPPPPPPPPPP 210

RESULT 3
US-11-036-256-96

/ Sequence 96, Application US/11036256
/ Publication No. US20060026719A1
/ GENERAL INFORMATION:
/ APPLICANT: KIELISZEWSKI, MARCIA
/ APPLICANT: XU, JIANPENG
/ TITLE OF INVENTION: METHODS OF PRODUCING PEPTIDES/PROTEINS IN PLANTS AND
/ FILE REFERENCES: 27211/04130
/ CURRENT APPLICATION NUMBER: US/11/036,256
/ PRIOR FILING DATE: 2005-01-14
/ PRIOR APPLICATION NUMBER: 60/602,562
/ PRIOR FILING DATE: 2004-08-18
/ PRIOR APPLICATION NUMBER: 60/582,027
/ PRIOR FILING DATE: 2004-06-22
/ PRIOR APPLICATION NUMBER: 60/536,486
/ PRIOR FILING DATE: 2004-01-14
/ NUMBER OF SEQ ID NOS: 173
/ SOFTWARE: PatentIn Ver. 3.3
/ SEQ ID NO 96
/ LENGTH: 1001
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence: Synthetic
/ OTHER INFORMATION: amino acid motif
/ NAME/KEY: MOD_RES
/ LOCATION: (1001)
/ OTHER INFORMATION: Lys, Ser, Thr, Ala, Gly, Val, or any amino acid
/ FEATURE:
/ OTHER INFORMATION: This sequence may encompass 2-1,000 residues of Pro
/ OTHER INFORMATION: between positions 1 and 1,000
US-11-036-256-96

Query Match 12.5%; Score 444; DB 7; Length 1001;
Best Local Similarity 35.5%; Pred. No. 4,4e-18;
Matches 94; Conservative 0; Mismatches 115; Indels 56; Gaps 2;
QY 34 PPAPASANI.PKPNQTPPTSDLSERFVSAPPPPPPPPPPTPMPPIAGSPSPSPA 93
Db 1 PPP 48
QY 94 ASKPPTPMPPIAGPBEAPKPPPTPMPPIAGPBEAPKPPPTPMPPIAGPAPPTESQLAPP 153
Db 49 PPP 108
QY 154 RPPPTGTPGAGQQPSSPAHVSHGPHQRRTAPAPMAKMPIGSPSPSPSPSPA 213
Db 109 PPP 168
QY 214 PPTRPAPQHSRRARRGHRYTDTERNVKGATGTSIQARLRABEASGAQLAPGTBSPAP 273
Db 169 PPPPPPP-----PPPPPPPP 184
QY 274 LQGRSYLAPTRPAPTEPPSPSP 298
Db 185 PPPPPPPPPPPPPPPPPPPPPPP 209

RESULT 4
US-11-096-568A-28799
/ Sequence 28799, Application US/11096568A
/ Publication No. US20060048240A1
/ GENERAL INFORMATION:
/ APPLICANT: Alexandrov, Nikolai et al.
/ TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
/ FILE REFERENCES: 2750-1592PUS2
/ CURRENT APPLICATION NUMBER: US/11/096,568A
/ PRIOR FILING DATE: 2005-04-01
/ NUMBER OF SEQ ID NOS: 34471
/ SEQ ID NO 28799
/ LENGTH: 717


```
/ Sequence 89, Application US/11036256
/ Publication No. US20060026719A1
/ GENERAL INFORMATION:
/ APPLICANT: KIELISZEWSKI, MARCIA
/ TITLE OF INVENTION: METHODS OF PRODUCING PEPTIDES/PROTEINS IN PLANTS AND
/ FILE REFERENCE: 27211/04130
/ CURRENT APPLICATION NUMBER: US/11/036,256
/ PRIOR FILING DATE: 2005-01-14
/ PRIOR APPLICATION NUMBER: 60/602,562
/ PRIOR FILING DATE: 2004-08-18
/ PRIOR APPLICATION NUMBER: 60/582,027
/ PRIOR FILING DATE: 2004-06-22
/ PRIOR APPLICATION NUMBER: 60/536,486
/ PRIOR FILING DATE: 2004-01-14
/ NUMBER OF SEQ ID NOS: 173
/ SOFTWARE: PatentIn Ver. 3.3
/ SEQ ID NO 89
/ LENGTH: 411
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence: Synthetic
/ OTHER INFORMATION: amino acid construct
US-11-036-256-89
```

```
Query Match      10.3%; Score 366.5; DB 7; Length 411;
Best Local Similarity 32.0%; Pred. No. 4.4e-14;
Matches 112; Conservative 18; Mismatches 131; Indels 89; Gaps 18;

QY 11 PHEGMAPDMAQAPFDPSPASPPRPAASANLPKNGQTPPT-----SDLSERV 62
DB 79 PPLGHSF-----PPLTSPSPPTPPPGHSPPLSPSPPTPPPLGHSPPPLTSPPT 133
QY 63 SAPPP-----PPPPPPPPPPPPM-PIAGBPP-----SPBPAKSPPTPPMPLAGPPAPKPP 114
DB 134 PTPPPGPHSPPLSPSPPTPPPLGHSPPPLSPSPPTPPPGVGPSPPLSPSPPT 193
QY 115 PTPPP-----PIAGBPPAPKPPPT--PWIAGPAPTPSPQLAPPP-----PTPQ 159
DB 194 PTPPPGPHSPPLSPSPPTPPPGHSPPLSPSPPTPPPLGHSPPPLSPSPPTPP 252
QY 160 TPTG--APQPPSPAP--HVPSHGPHQPPRT-----APAPWAKAPIGBPP 202
DB 253 PPGGHSPPPLSPSPPTPPPLGHSPPPLSPSPPTPPPGHSPPLSPSPPTPPPL 312
QY 203 APGR--PSASPPAPPTPPAPQHSRRARGRHRTDTERNVGKATGPSIOAKLAEASG 260
DB 313 GPHSPPLSPSPPTPPPPGPHS-----PPPLSPSPPTPPPLG 351
QY 261 AQLAPGT--BSPAPLQGPSYL--APTPAPTEPPP-----SPSP 298
DB 352 PHSPPPLSPSPPTPPPGHSPPLSPSPPTPPPLGHSPPPLSPSP 401
```

```
RESULT 7
US-11-096-568A-2501
/ Sequence 2501, Application US/11096568A
/ Publication No. US20060048240A1
/ GENERAL INFORMATION:
/ APPLICANT: Alexandrov, Nikolai et al.
/ TITLE OF INVENTION: Sequence-determined DNA Fragments and Corresponding Polypeptides
/ FILE REFERENCE: 2750-1582PUS2
/ CURRENT APPLICATION NUMBER: US/11/096,568A
/ NUMBER OF SEQ ID NOS: 34471
/ SEQ ID NO 2501
/ LENGTH: 681
/ TYPE: PRT
/ ORGANISM: Arabidopsis thaliana
/ FEATURE:
```

```
/ NAME/KEY: misc_feature
/ LOCATION: (1)..(681)
/ OTHER INFORMATION: Cereus Seq. ID no. 14308949
US-11-096-568A-2501
```

```
Query Match      10.1%; Score 359.5; DB 7; Length 681;
Best Local Similarity 24.6%; Pred. No. 1.9e-13;
Matches 142; Conservative 70; Mismatches 207; Indels 159; Gaps 26;

QY 34 PPAPASANLPKNGQTPPTSDLSERFVSAPPPPPPPPTPPMPLAGBPP----- 88
DB 12 PPSNSSTTAPPLQTPPT-----PSAPPPVTPPPSPQSBPPVSSPPPPVVS 62
QY 89 SPEPAKSPPTPPMPLAGBPPAPKPPTPPMBIAGEPA-----PPKPP-----TPMPPIA 139
DB 63 SPSSSPSPSPV--ITSPPTVASPPPPVVIASPPPTPATTPPAPQTVSPPPPPA 121
QY 140 GAPPTPESQLAPPPPTQTPTGAPQD--SEPAHVPHSHGPHQPPRTAPAPMAKMPIG 198
DB 122 SPSP--PAPTTNPPPKSPSPGKTPSPGKTPSPKPPSPPTP--TTTSP----- 172
QY 199 BPPAPSR--PSAPAP-----PTTPAPQHSRRARGRHRTDTERNVGKATGPSIOA 251
DB 173 -PPPATSPSPSPNPTDESTLAPPPPLPVR-----EKPIAK-PTGPA--- 215
QY 252 RLRAEASGAQAPGTPSPAPLG-----PPSYLAPT 285
DB 216 -----SNNGNTLSPSSPGKSEVGTGIVAIGVIVLPSLFWGVWFTPKRKRPCT 270
QY 286 RAPTPTPPS--SPQR-----NSGRARRRVHPLDAAQHAQAQDSTIAATGGRRRR 338
DB 271 FVGTPMPAPAYSPQSDVVLFNSSSAPPKRSHSGSYMYASSDS-----GNVSNQSRKF 327
QY 339 AAPDLDA-----TOKSLRPAAGPKVK--VAPQKATKPKVVSQGRMRM-----VHAL 388
DB 328 SYDELQVTSGRSEKTLDEGGFGCVYKGVLSDEGVAVKQKIGSGQSERFKAEVEII 387
QY 389 TRIN-----LGLSDEKTELDLHARVRNRPGSYQIADVGLKQAGKTTLLAIGSTL 441
DB 388 SVHHRHLVTLVGYCISEQHRLLVYDVNNTL--HYHLAPRPVMTWETRVVAAGAR 446
QY 442 A-----QVADRII-----ADDPGAGNLADRVGRGAGATTAADVLAEK 480
DB 447 GIAYLHEDCPRIIHRDIKSSNILLDNSEFALVADFGAKIAQEL----- 491
QY 481 ELSHYNDIRAHTSVNAVNLVLPAPPEYSAQALSDAD 518
DB 492 -----DLNTHVSTRVMGTGMADEYATSGLSKAD 523
```

```
RESULT 8
US-11-036-256-33
/ Sequence 33, Application US/11036256
/ Publication No. US20060026719A1
/ GENERAL INFORMATION:
/ APPLICANT: KIELISZEWSKI, MARCIA
/ TITLE OF INVENTION: METHODS OF PRODUCING PEPTIDES/PROTEINS IN PLANTS AND
/ FILE REFERENCE: 27211/04130
/ CURRENT APPLICATION NUMBER: US/11/036,256
/ PRIOR FILING DATE: 2005-01-14
/ PRIOR APPLICATION NUMBER: 60/602,562
/ PRIOR FILING DATE: 2004-08-18
/ PRIOR APPLICATION NUMBER: 60/582,027
/ PRIOR FILING DATE: 2004-06-22
/ PRIOR APPLICATION NUMBER: 60/536,486
/ PRIOR FILING DATE: 2004-01-14
/ NUMBER OF SEQ ID NOS: 173
/ SOFTWARE: PatentIn Ver. 3.3
/ SEQ ID NO 33
/ LENGTH: 346
/ TYPE: PRT
```

```

; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: amino acid construct
US-11-036-256-33

Query Match
  9.8%; Score 348; DB 7; Length 346;
  Best Local Similarity 33.9%; Pred. No. 4,2e-13;
  Matches 101; Conservative 22; Mismatches 117; Indels 58; Gaps 17;

QY 29 PSASPP-APASANLPKNGQTPPTSDLSERFVSAPP-PPPPPPPTPM-PIA 83
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 26 PHSPPPLSPSPPTPPPLGPHSPPT-LSPPPTPPPGHSPPLSPSPPTPPPLG 82
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 84 AGSP-SPSPASKPP--TPPMPIA-GPBAPPKPPTPPMPIAGPBAPKPTP- 134
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 83 PHSPPPLSPSPPTPPPGHSPPLSPSPPTPPPLGPHSPPTPLSPSTPTPPPGPHS 142
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 135 PMDIAPAPPTESQIAPPP-PPPTPTG-APQPSAP-HVPSGHQPR 184
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 143 PPSPLSPSPPT-PLGPHSPPTLSPSPPTPPPGHSPPLSPSPPTPPPLGPH- 198
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 185 RTAPAPPMAMPKGPBPAPSR-PSASPPPTTPAPQHSRARKGHRVTETENVK 242
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 199 --SPPTLSPSPPTPPPGHSPPLSPSPPTPPPLGPHS----- 237
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 243 VATGPIQARLRBBSGAQIAPOTBSPAPLQPSYAP-PTRAPPT-EPPSPSP 298
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 238 --PPPTLSPSPPTPPPGHSPPTLTPPLAPPPPLPHGTPPLTPPLAPPT 293
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 9
US-11-169-232-14
; Sequence 14, Application US/11169232
; Publication No. US20060025570A1
; GENERAL INFORMATION:
; APPLICANT: Turner, Katherine
; Clark, Stephen C.
; Jacobs, Kenneth
; Hewick, Rodney M.
; Geener, Thomas G.
; TITLE OF INVENTION: Megakaryocyte Stimulating Factors
; NUMBER OF SEQUENCES: 143
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Genetics Institute, Inc.
; STREET: 87 Cambridgepark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/11/169,232
; FILING DATE: 28-Jun-2005
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/10/124,557
; FILING DATE: 16-Apr-2002
; APPLICATION NUMBER: US 07/643,502
; FILING DATE: 18-Jan-1991
; APPLICATION NUMBER: US 07/546,114
; FILING DATE: 29-Jun-1990
; APPLICATION NUMBER: US 07/457,196
; FILING DATE: 29-Dec-1989
; APPLICATION NUMBER: US 07/390,901
; FILING DATE: 08-Aug-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Geert, Luann
; REGISTRATION NUMBER: 31,822
```

```

; REFERENCE/DOCKET NUMBER: GI 5190
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 876-1170
; TELEFAX: (617) 876-5851
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 941 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 14:
US-11-169-232-14

Query Match
  9.8%; Score 347; DB 7; Length 941;
  Best Local Similarity 25.4%; Pred. No. 1.4e-12;
  Matches 146; Conservative 58; Mismatches 259; Indels 112; Gaps 29;

QY 28 DPASPP-APASANLPKNGQTPPTSDLSERFVSAPP-PPPPPTPM 80
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 251 EPAPTPPKBPTTPPKBAPPTTKBAPPTP--KEBAPAPKBPAPTPPKBAP 308
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 81 PIAGSP-----PSPPASKPTP-PMPIAGBPAP-PPK 114
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 309 PTTKSPSPPTPKBAPPTTKBAPPTTKBAPPTTKBAPPTTKBAPPTTKBAP 368
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 115 PTP-----PMPIAGBPAP-PPKPTP-----PMPIA-GPAP-TPESQIAPPPPTPQ 159
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 369 PAPTTPKKAAPTTPKBPAPPTPKBAPPTTKBAPAPAPAPAPAPAPAPAPAP 428
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 160 TPTG-APQPSAPHPVPSH---GPHQPRTPAPPMAMPKGPBPAPSR-ASPA 213
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 429 TPEKLAPTTPKBPAPPTPEBAPPTPEBAPPTPEBAPPTPEBAPPTPEBAP 488
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 214 P-----PTRAPQHSRARKGHRVTETENVGVAAGPSIQAIRLBBSGAQIAPGT- 267
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 489 PAPTTPKBPAPPTPK-----ETAPTPKGAPTTLKEBAPPTPKBPAPPTPK 540
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 268 EPS-----PAPLQPSYAPPTTPAPPTP-PPSPQSGNSGRARRVHPDLAAQHA 320
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 541 EPSTTSQKBPAPPTTPKBPAPPTPKBAPPTTPKBPAPPTTPKBPAPPTPK 599
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 321 AOPDSITAAITGGRRRKRAAPDLATQSLRPAKPK-VKVKPQPKAT---KPPKV 376
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 600 --PKELAPTTKGPSTSDKBPAPTPPKETAPPT--PKBAPPTPKBPAPPT 655
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 377 SCQGMHWHAALTRINLGSPEKYLDAHAR-----VARNPQSGQIANGKAGAKT 431
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 656 SEVSTPTTKBPTTIH--KSPBSTP-ELSAEPTPALNSPK-----SPGVPT 702
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 432 TITPAIGSTLAQVRADRIALD--ADPGAAGNLADRVGROSGATIDVLAKEKLSHYNDIR 489
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 703 KTPPAATKPEKTTAAKDKTTRBDLTTPETTTAPAPKATKATITTKTSKITATTQV 762
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 490 AHTSVNAVILVLP-----APEYSAQALSDAD 518
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 763 STTQDTPPKITTLTKTLAPKVTTKKITTTT 797
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 10
US-11-169-232-84
; Sequence 84, Application US/11169232
; Publication No. US20060025570A1
; GENERAL INFORMATION:
; APPLICANT: Turner, Katherine
; Clark, Stephen C.
; Jacobs, Kenneth
; Hewick, Rodney M.
; Geener, Thomas G.
; TITLE OF INVENTION: Megakaryocyte Stimulating Factors
; NUMBER OF SEQUENCES: 143
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Genetics Institute, Inc.
; STREET: 87 Cambridgepark Drive
```


Db 316 EPAATTPKEPTTPPKBAPATTKBAPATTP--KBAPATPAKKAAPTTKBPAPTTKBPBA 373
Qy 81 PIAAGEP-----PSBPAAKSPPTP--PMPIAGEBPAP--PKP 114
Db 374 PTTTKBSPPTTKBPAPATTTKSAATTTKBPAPATTTKBPAPATTTKBPAPATTTKBPAPATTTKBP 433
Qy 115 PTP-----PMPIAGEBPAP--PKPPTP-----PMPIA--GPAP--TPPESQALAPRPPTPQ 159
Db 434 PAPTTPKBPAPATTTKBPAPATTTKBPAPATTTKBPAPATTTKBPAPATTTKBPAPATTTKBPAPATTTKBP 493
Qy 160 TPTG--APQBPSPAPHPHSH-----GHPQPRRTAPAPWAKMPIGBPAPBPAPBP--ASPAE 213
Db 494 TPKKLAATTPKBPAPATTTKBPAPATTTKBPAPATTTKBPAPATTTKBPAPATTTKBPAPATTTKBPAPATTTKBP 553
Qy 214 P-----PTPAPQHSRRARRGHRVYTDTERNVGVATGPSIOARLAEASGAOLAPGT-- 267
Db 554 PAPTTPKBPAPATTPK-----ETAPTTPKGTAPTTLKBPAPATTTKBPAPATTTKBPAPATTTKBPAPATTTKBP 605
Qy 268 EPS-----PAPLQGPSTYAPPTTPAPPTBP--PPSPSPQNSGRRAERVRHPLAQAHA 320
Db 606 EPTSTSDKPAFTTPKGTAPTTKBPAPATTTKBPAPATTTKBPAPATTTKBPAPATTTKBPAPATTTKBPAPATTTKBP 664
Qy 321 AOPDSITAAATGGRRRRKAPDLDATQKSLRPAKGPK--VKVYKQPKAT--KPPKV 376
Db 665 --PKELAPTTTKGPTSTSDKAPATTTKBPATTTKBPATTTKBPATTTKBPATTTKBPATTTKBPATTTKBPATTTKBP 720
Qy 377 SQRGWHVVALTRINIGLSPDEKYEELDLHAR-----VRNPPGSGYQIAYVGLKGAGKT 431
Db 721 SVSTPTTKGPTTIH--KSPDSTP--ELSAEPTPKLSPK-----EPGPTT 767
Qy 432 TLTAALGSTLAQVRADRIALD--ADPGANLADRVGROGATITADVLAKESLHNDIR 489
Db 768 KTPAAATPEMTTAKDKTERDLRTTBETTTAAPKMKETATTTKTESKITATTTQVT 827
Qy 490 AHTSVANVLEVP-----APEYSSAORALSDAD 518
Db 828 STTQDTPPKITTKTTTLAATKVTTKKITTTT 862

RESULT 12
US-11-169-232-58
; Sequence 58, Application US/11169232
; Publication No. US20060025570A1
; GENERAL INFORMATION:
; APPLICANT: Turner, Katherine
; Clark, Stephen C.
; Jacobs, Kenneth
; Hewick, Rodney M.
; Gessner, Thomas G.
; TITLE OF INVENTION: Megakaryocyte Stimulating Factors
; NUMBER OF SEQUENCES: 143
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Genetics Institute, Inc.
; STREET: 87 Cambridgepark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/11/169,232
; FILING DATE: 28-Jun-2005
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/10/124,557
; FILING DATE: 16-Apr-2002
; APPLICATION NUMBER: US 07/643,502
; FILING DATE: 18-Jan-1991
; APPLICATION NUMBER: US 07/546,114

; FILING DATE: 29-JUN-1990
; APPLICATION NUMBER: US 07/457,196
; FILING DATE: 29-DEC-1989
; APPLICATION NUMBER: US 07/390,901
; FILING DATE: 08-AUG-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Cseri, Luanm
; REGISTRATION NUMBER: 31,822
; REFERENCE/DOCKET NUMBER: GI 5190
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 876-1170
; TELEFAX: (617) 876-5851
; INFORMATION FOR SEQ ID NO: 58:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1049 amino acids
; TYPE: amino acid
; TOPOLOGY: Linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 58:
US-11-169-232-58
Query Match 9.84; Score 347; DB 7; Length 1049;
Best Local Similarity 25.44; Pred. No. 1.6e-12;
Matches 146; Conservative 58; Mismatches 259; Indels 112; Gaps 29;
Qy 28 DPASAFP--PAPASAMLPKENGQTPPTSDLSRFVSAAPPPPP-----PPPPPPPTPM 80
Db 359 EPAATTPKEPTTPPKBAPATTTKBPAPATTTKBPAPATTTKBPAPATTTKBPAPATTTKBPAPATTTKBPAPATTTKBP 416
Qy 81 PIAAGEP-----PSBPAAKSPPTP--PMPIAGEBPAP--PKP 114
Db 417 PTTTKBSPPTTKBPAPATTTKBPAPATTTKBPAPATTTKBPAPATTTKBPAPATTTKBPAPATTTKBPAPATTTKBP 476
Qy 115 PTP-----PMPIAGEBPAP--PKPPTP-----PMPIA--GPAP--TPPESQALAPRPPTPQ 159
Db 477 PAPTTPKBPAPATTTKBPAPATTTKBPAPATTTKBPAPATTTKBPAPATTTKBPAPATTTKBPAPATTTKBPAPATTTKBP 536
Qy 160 TPTG--APQBPSPAPHPHSH-----GHPQPRRTAPAPWAKMPIGBPAPBPAPBP--ASPAE 213
Db 537 TPKKLAATTPKBPAPATTTKBPAPATTTKBPAPATTTKBPAPATTTKBPAPATTTKBPAPATTTKBPAPATTTKBPAPATTTKBP 596
Qy 214 P-----PTPAPQHSRRARRGHRVYTDTERNVGVATGPSIOARLAEASGAOLAPGT-- 267
Db 597 PAPTTPKBPAPATTPK-----ETAPTTPKGTAPTTLKBPAPATTTKBPAPATTTKBPAPATTTKBPAPATTTKBPAPATTTKBP 648
Qy 268 EPS-----PAPLQGPSTYAPPTTPAPPTBP--PPSPSPQNSGRRAERVRHPLAQAHA 320
Db 649 EPTSTSDKPAFTTPKGTAPTTKBPAPATTTKBPAPATTTKBPAPATTTKBPAPATTTKBPAPATTTKBPAPATTTKBPAPATTTKBP 707
Qy 321 AOPDSITAAATGGRRRRKAPDLDATQKSLRPAKGPK--VKVYKQPKAT--KPPKV 376
Db 708 --PKELAPTTTKGPTSTSDKAPATTTKBPAPATTTKBPAPATTTKBPAPATTTKBPAPATTTKBPAPATTTKBPAPATTTKBP 763
Qy 377 SQRGWHVVALTRINIGLSPDEKYEELDLHAR-----VRNPPGSGYQIAYVGLKGAGKT 431
Db 764 SVSTPTTKGPTTIH--KSPDSTP--ELSAEPTPKLSPK-----EPGPTT 810
Qy 432 TLTAALGSTLAQVRADRIALD--ADPGANLADRVGROGATITADVLAKESLHNDIR 489
Db 811 KTPAAATPEMTTAKDKTERDLRTTBETTTAAPKMKETATTTKTESKITATTTQVT 870
Qy 490 AHTSVANVLEVP-----APEYSSAORALSDAD 518
Db 871 STTQDTPPKITTKTTTLAATKVTTKKITTTT 905

RESULT 13
US-11-169-232-104
; Sequence 104, Application US/11169232
; Publication No. US20060025570A1
; GENERAL INFORMATION:
; APPLICANT: Turner, Katherine
; Clark, Stephen C.

```

; Jacob, Kenneth
; Hewick, Rodney M.
; Geener, Thomas G.
; TITLE OF INVENTION: Megakaryocyte Stimulating Factors
; NUMBER OF SEQUENCES: 143
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 Cambridgepark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02140
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/11/169,232
; FILING DATE: 28-Jun-2005
; CLASSIFICATION: <Unknown>
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/10/124,557
; FILING DATE: 16-Apr-2002
; APPLICATION NUMBER: US 07/643,502
; FILING DATE: 18-JAN-1991
; APPLICATION NUMBER: US 07/546,114
; FILING DATE: 29-JUN-1990
; APPLICATION NUMBER: US 07/457,196
; FILING DATE: 29-DEC-1989
; APPLICATION NUMBER: US 07/390,901
; FILING DATE: 08-AUG-1989
;
; ATTORNEY/AGENT INFORMATION:
; NAME: Geert, Luann
; REGISTRATION NUMBER: 31,822
; REFERENCE/DOCKET NUMBER: GI 5190
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 876-1170
; TELEFAX: (617) 876-5851
;
; INFORMATION FOR SEQ ID NO: 104:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1140 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 104:
US-11-169-232-104

Query Match          9.8%; Score 347; DB 7; Length 1140;
Best Local Similarity 25.4%; Pred. No. 1.7e-12;
Matches 146; Conservative 58; Mismatches 259; Indels 112; Gaps 29;

QY 28 DASASP--PAPASALPKPNCGTPTPTSDLSERVSAAPPPPP-----PPPPPPPTPM 80
DB 450 EBPATPKKPTTPPKKPAATTKKPAATTP--KKKPAATKKAATTPPKKPAATTKKPA 507
QY 81 PIAAGRP-----PSPEPAASKPTP--PMPPIAGPEBAP--PKK 114
DB 508 PTTTKKPSPTTPKPAATTTKSAPTTTKKEAPTTTSAPPTPKKPSPTTTKKEAPTTPKK 567
QY 115 PTP-----PMPPIAGPEBAP--PKKPTP-----PMPPIA--GPAP--TTTESGLAPRPPTPQ 159
DB 568 PAPPTPKKPAATTPPKKPAATTPPKKPAATTKKPAATPKKPAATTPPKKPAATTPPKK 627
QY 160 TPTG--APQEPESAPHVPSH-----GPHQPRRTAPAPPMAMPGEPPPAASRPS--ASPAE 213
DB 628 TBEKLAPTTKKPAATTPBEELATTPBEBPTTPPEBPAPTTTAAAPNTPKKPAATTPKE 667
QY 214 P-----PTRPAPQHSRRARRGRHYRTDERVGVATGPSIOARLRABEASGAQLAGT-- 267
DB 688 PAPPTTKKPAATTPK-----ETAPTTKKGAPATTKKPAATTKKPAATTKKPAATTKK 739
QY 268 EBS-----PAPIGPQPSYLAAPTTPAPATP--PPSPSPQPSNGRRARRVHVDLAOHAA 320

```

```

DB 740 EPTSTSDKPAATTPPKKPAATTPPKKPAATTKKPAATTKKPAATTKKPAATTKKPA 798
QY 321 AOPDSITAAATGGRARRKRAAPLDAATQKSLRPAAGPK--VKKVPQKPAAT--KPKKV 376
DB 799 --PKELAPTTKGPSTTSDKPAATTPPKKPAATTP--PKKPAATTPPKKPAATTPETPPPT 854
QY 377 SQGRKRWHAALTRINLGLSPDEKHELDLHAR-----VRNRGSGQIAVGLKAGAGKT 431
DB 855 SEVSTPTTKKPTTH--KSPDSPT--ELSAEPTPKALENSPK-----EPGVPTT 901
QY 432 TLTAAGSTLAQVRADRIALD--ADPGAGNLADRVGROSQATVADVLAKEKLSHYNDIR 489
DB 902 KTPPAATPKKPTTAAKDTTERBDRITTPETTTAAPKATKATATTETTESKITATTQT 961
QY 490 AHTSVANVLEVLV-----APEYSAORALSDAD 518
DB 962 STTQDTTPFKITTLTKTLAPKVTTKKTIITTE 996

RESULT 14
US-11-169-232-44
; Sequence 44, Application US/11169232
; Publication No. US20060025570A1
;
; GENERAL INFORMATION:
; APPLICANT: Turner, Katherine
; Clark, Stephen C.
; Jacobs, Kenneth
; Hewick, Rodney M.
; Geener, Thomas G.
;
; TITLE OF INVENTION: Megakaryocyte Stimulating Factors
; NUMBER OF SEQUENCES: 143
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 Cambridgepark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02140
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/11/169,232
; FILING DATE: 28-Jun-2005
; CLASSIFICATION: <Unknown>
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/10/124,557
; FILING DATE: 16-Apr-2002
; APPLICATION NUMBER: US 07/643,502
; FILING DATE: 18-JAN-1991
; APPLICATION NUMBER: US 07/546,114
; FILING DATE: 29-JUN-1990
; APPLICATION NUMBER: US 07/457,196
; FILING DATE: 29-DEC-1989
; APPLICATION NUMBER: US 07/390,901
; FILING DATE: 08-AUG-1989
;
; ATTORNEY/AGENT INFORMATION:
; NAME: Geert, Luann
; REGISTRATION NUMBER: 31,822
; REFERENCE/DOCKET NUMBER: GI 5190
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 876-1170
; TELEFAX: (617) 876-5851
;
; INFORMATION FOR SEQ ID NO: 44:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1270 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 44:

```


Tue Apr 18 07:09:49 2006

us-10-620-246-70.rapbn

Page 10

Search completed: April 14, 2006, 18:42:14
Job time : 31.5386 Secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using SW model

Run on: April 14, 2006, 17:03:40 ; Search time 37.9711 Seconds
(without alignments)
1133.996 Million cell updates/sec

Title: US-10-620-246-88

Perfect score: 479
Sequence: 1 MEKMSHDPADIGTQVSDN.....VQDVARTYSQIDGAGVFA 98

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: A_Geneseq_21:*
2: geneseqp1980s:*
3: geneseqp1990s:*
4: geneseqp2000s:*
5: geneseqp2001s:*
6: geneseqp2002s:*
7: geneseqp2003s:*
8: geneseqp2004s:*
9: geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	479	100.0	98	2	AAW72927 Mycobacte
2	479	100.0	98	2	AAW72927 Mycobacte
3	479	100.0	99	4	AAW72927 Mycobacte
4	479	100.0	99	4	AAW72927 Mycobacte
5	102.5	21.4	102	5	AAU70949 M. tuberc
6	102.5	21.4	102	5	AAU70949 M. tuberc
7	102.5	21.4	102	5	AAU70949 M. tuberc
8	92.5	19.3	639	6	ABU36685 M. tuberc
9	92.5	19.3	639	6	ABU36685 M. tuberc
10	90	18.8	588	6	ABU36685 M. tuberc
11	90	18.8	588	6	ABU36685 M. tuberc
12	87.5	18.3	606	6	ABU36685 M. tuberc
13	87.5	18.3	606	6	ABU36685 M. tuberc
14	87	18.2	102	5	AAU70953 M. tuberc
15	86.5	18.1	783	9	ABU36685 M. tuberc
16	86.5	18.1	783	9	ABU36685 M. tuberc
17	85.5	17.8	543	9	ABU36685 M. tuberc
18	85.5	17.8	543	9	ABU36685 M. tuberc
19	85	17.7	487	6	ABU36685 M. tuberc
20	83	17.3	576	6	ABU36685 M. tuberc
21	83	17.3	576	6	ABU36685 M. tuberc
22	82	17.1	562	6	ABU36685 M. tuberc
23	81	16.9	1011	6	ABU36685 M. tuberc
24	80.5	16.8	319	6	ABU36685 M. tuberc

25	80.5	16.8	853	9	ABU36682 Protein e
26	80.5	16.8	914	6	ABU36682 Protein e
27	80.5	16.8	914	6	ABU36682 Protein e
28	80.5	16.8	1660	9	ABU36682 Protein e
29	80	16.7	588	6	ABU36682 Protein e
30	80	16.7	767	9	ABU36682 Protein e
31	80	16.7	923	6	ABU36682 Protein e
32	80	16.7	923	6	ABU36682 Protein e
33	79.5	16.6	137	5	AAU70987 M. tuberc
34	79.5	16.6	370	5	AAU70987 M. tuberc
35	78.5	16.4	957	6	ABU36682 Protein e
36	78.5	16.4	102	5	ABU36682 Protein e
37	77.5	16.2	603	6	ABU36682 Protein e
38	76.5	16.0	439	6	ABU36682 Protein e
39	75.5	15.8	100	5	ABU36682 Protein e
40	75.5	15.8	464	8	ADN25249 Bacteri
41	75.5	15.8	505	6	ABU36682 Protein e
42	75.5	15.8	558	7	ADG15732 M. tuberc
43	75.5	15.8	615	6	ABU36682 Protein e
44	75.5	15.8	615	6	ABU36682 Protein e
45	75.5	15.8	1306	6	ABU36682 Protein e

ALIGNMENTS

RESULT 1	AAW72927	standard; protein; 98 AA.
ID	AAW72927	
XX	AAW72927	
AC	AAW72927	
DT	21-JAN-1999	(first entry)
XX		
DB	Mycobacterium tuberculosis antigen RD1-ORF3.	
XX		
KW	Mycobacterium tuberculosis; antigen; vaccine; immunological; immunogen; infection.	
XX		
OS	Mycobacterium tuberculosis.	
XX		
PN	W09844119-A1.	
XX		
PD	08-OCT-1998.	
XX		
PR	01-APR-1998; 98MO-DK000132.	
XX		
PR	02-APR-1997; 97DK-00000376.	
PR	18-APR-1997; 97US-0044624P.	
PR	10-NOV-1997; 97DK-00001277.	
PR	05-JAN-1998; 98US-0070488P.	
XX		
PA	(STAT-) STATENS SERUM INST.	
XX		
PI	Andersen P, Nielsen R, Rosenkrands I, Welling K, Rasmussen PB, Oettinger T, Florio W;	
XX		
DR	WPI: 1998-542705/46.	
XX		
PT	New isolated mycobacteria polypeptides and nucleic acids - used for developing products for the diagnosis of or vaccination against mycobacterial infections, particularly tuberculosis.	
XX		
PS	Claim 1, Page 196, 163pp; English.	
XX		
CC	The present sequence represents a Mycobacterium tuberculosis protein. Products from the present invention, which describes protein fragments and nucleic acid fragments derived from M. tuberculosis, can be used in the detection of and prevention of mycobacterial infections. In particular, the proteins and nucleic acids can be used for the diagnosis of or vaccination against tuberculosis caused by M. tuberculosis, M. africanum or M. bovis	
CC		

```

XX      Sequence 98 AA;
SQ
Query Match      100.0%; Score 479; DB 2; Length 98;
Best Local Similarity 100.0%; Pred. No. 3.3e-46;
Matches 98; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MEMSHDPIADIGTGVSDNALHGVTAAGTALTSTVGLVPAGADEVSAQAATAFTSEGIQ 60
DB      1 MEMSHDPIADIGTGVSDNALHGVTAAGTALTSTVGLVPAGADEVSAQAATAFTSEGIQ 60

QY      61 LLASNASADQLHRAGEAVODVARTYSQIDPGAAGVFA 98
DB      61 LLASNASADQLHRAGEAVODVARTYSQIDPGAAGVFA 98

RESULT 2
AAV21944 standard; protein; 98 AA.
XX
AC      AAV21944;
XX
DT      06-SEP-1999 (first entry)
XX
DE      Amino acid sequence of antigen RD1-ORF3.
XX
KW      Immunogenic; Mycobacterium tuberculosis; immune response; infection;
KW      tuberculosis; fusion polypeptide; T-cell epitope; ESAT-6; MPT59; TB;
KW      pharmaceutical; vaccination; M. africanum; M. bovis; CFP7A; CFP30A;
KW      CFP7B; CFP13; CFP27; CFP30A; RD1-ORF; CFP10A; CFP16; CFP19; CFP23;
KW      CFP25A; CFP30B; CFP7B.
XX
OS      Mycobacterium tuberculosis.
XX
PN      WO924577-A1.
XX
PD      20-MAY-1999.
XX
PF      08-OCT-1998; 98WO-DK000438.
XX
PR      10-NOV-1997; 97DK-00001277.
PR      05-JAN-1998; 98US-0070468P.
PR      01-APR-1998; 98WO-DK000132.
XX
PA      (STAT-) STATENS SERUM INST.
PI      Andersen P, Skjot R;
XX
DR      WPI; 1999-347282/29.
DR      N-PSDB; AAX81044.
XX
PT      New immunogenic fragment of Mycobacterium tuberculosis.
XX
PS      Example 2; Page 215; 265pp; English.
XX
CC      The invention describes a substantially pure immunogenic polypeptide
CC      fragment (I) from Mycobacterium tuberculosis that is able to evoke a
CC      protective immune response against infections by mycobacteria belonging
CC      to the tuberculosis complex. The invention provides a (1) fusion
CC      polypeptide comprising at least one polypeptide fragment (1) and at least
CC      one fusion partner; (2) a fusion polypeptide fragment comprising a T-cell
CC      epitope from M. tuberculosis protein ESAT-6, or MPT59 and a second
CC      different amino acid sequence from M. tuberculosis, and/or including a
CC      sequence which protects the first amino acid sequence from in vivo
CC      degradation or post-translational processing; (3) a nucleic acid fragment
CC      that encodes the above polypeptides. The polypeptides and nucleic acid
CC      are useful as pharmaceuticals, for diagnosis of and as antigens for
CC      vaccination against TB caused by Mycobacterium tuberculosis, africanum or
CC      bovis. The polypeptides are also useful for diagnosing ongoing or
CC      previous sensitization in an animal with bacteria belonging to the
CC      tuberculosis complex. The invention also describes the use of CFP7A or
CC      CFP30A or a T-cell epitope of for the induction of a strong immune
CC      response in a mammal; use of CFP7B, CFP19 or MPT59-ESAT6 or a T-cell

```

```

CC      epitope of for diagnosis of TB in a mammal by performing a DTH type skin
CC      test; use of CFP27, CFP30A, RD1-ORF2, RD1-ORF3, RD1-ORF5, MPT59-ESAT6,
CC      ESAT6-MPT59, CFP10A, CFP16, CFP19, CFP23, CFP25A, CFP30B, CFP7B or a T-
CC      cell epitope of for the preparation of an immunological composition; and
CC      for the preparation of a subunit vaccine
XX
SQ      Sequence 98 AA;
XX
Query Match      100.0%; Score 479; DB 2; Length 98;
Best Local Similarity 100.0%; Pred. No. 3.3e-46;
Matches 98; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MEMSHDPIADIGTGVSDNALHGVTAAGTALTSTVGLVPAGADEVSAQAATAFTSEGIQ 60
DB      1 MEMSHDPIADIGTGVSDNALHGVTAAGTALTSTVGLVPAGADEVSAQAATAFTSEGIQ 60

QY      61 LLASNASADQLHRAGEAVODVARTYSQIDPGAAGVFA 98
DB      61 LLASNASADQLHRAGEAVODVARTYSQIDPGAAGVFA 98

RESULT 3
AAB19843 standard; protein; 99 AA.
XX
AC      AAB19843;
XX
DT      05-MAR-2001 (first entry)
XX
DE      Mycobacterium tuberculosis protein MTBN2.
XX
KW      MTBN2; tuberculosis; BCG; vaccine; infection; diagnosis.
XX
OS      Mycobacterium tuberculosis.
XX
PN      WO20006157-A1.
XX
PD      09-NOV-2000.
XX
PF      04-MAY-2000; 2000WO-US012257.
XX
PR      04-MAY-1999; 99US-0132505P.
XX
PA      (PUBL-) PUBLIC HEALTH RES INST NEW YORK.
PI      Genaro ML;
XX
DR      WPI; 2001-007153/01.
DR      N-PSDB; AAB89036.
XX
PT      Novel polypeptide encoded by open reading frames present in Mycobacterium
PT      tuberculosis genome and not by the BCG strain of M. bovis, useful as
PT      vaccine and for diagnosing tuberculosis infection.
XX
PS      Claim 11; Fig 1; 35pp; English.
XX
CC      The present sequence is that of the Mycobacterium tuberculosis MTBN2
CC      protein. This is 1 of 8 proteins, i.e. MTBN1-8 (see AAB19842-49), encoded
CC      by 8 open reading frames (see AAB89035-42) identified as being present in
CC      the genome of M. tuberculosis but absent from the genome of the BCG
CC      strain of Mycobacterium bovis. MTBN1-8 represent reagents that are useful
CC      in discriminating between M. tuberculosis and BCG and, in particular, for
CC      diagnostic methods which discriminate between exposure of a subject to M.
CC      tuberculosis and vaccination with BCG. The invention features these MTBN
CC      polypeptides, functional fragments of them, DNA encoding them, vectors,
CC      transformed cells, and diagnostic, therapeutic, and prophylactic
CC      (vaccine) methods, including genetic vaccination methods
XX
SQ      Sequence 99 AA;

```

Query Match 100.0%; Score 479; DB 4; Length 99;
 Best Local Similarity 100.0%; Pred. No. 3.4e-46;
 Matches 98; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEKSHDPIADIGTGVSDNALHGVTAGSTALTSTVGLVPGADSVSAQAATAFTSSGIG 60
 |||||||
 DB 1 MEKSHDPIADIGTGVSDNALHGVTAGSTALTSTVGLVPGADSVSAQAATAFTSSGIG 60

QY 61 LLASNASADQDLHRAGBAVDVARTYSQIDGAGVFA 98
 |||||||
 DB 61 LLASNASADQDLHRAGBAVDVARTYSQIDGAGVFA 98

RESULT 4
 ADG15730 standard; protein; 99 AA.
 ADG15730
 XX
 AC ADG15730;
 XX
 DT 26-FEB-2004 (first entry)
 XX
 DE M tuberculosis Rv3872 protein for generating peptide epitopes.
 XX
 KM antibacterial; antitubercular; tuberculostatic; vaccine;
 KM diagnostic reagent; epitope; RD1 region; RD2 region;
 KM Mycobacterium tuberculosis; Mycobacterium bovis; Mycobacterium africanum.
 OS Mycobacterium tuberculosis.
 XX
 PN W02003093307-A2.
 XX
 PD 13-NOV-2003.
 XX
 PF 28-APR-2003; 2003WO-GB001815.
 XX
 PR 27-APR-2002; 2002GB-00009723.
 PR 27-APR-2002; 2002GB-00009724.
 XX
 PA (UKEN-) UK SEC FOR ENVIRONMENT FOOD & RURAL AFF.
 XX
 PI Cockle PJ, Vordermeier HM, Gordon SV, Hewinson RG;
 XX
 DR WPI; 2003-903652/82.
 XX
 PT New diagnostic reagents comprising a peptide having an epitope from
 PT polypeptides Rv1985, Rv3878, Rv1983, Rv3873 or Rv3879 derived from RD1
 PT and RD2 regions of Mycobacterium, useful as vaccines against
 PT Mycobacterium infections.
 XX
 PS Claim 19; SEQ ID NO 2; 77pp; English.
 XX
 CC The invention relates to a novel diagnostic reagent comprising a peptide
 CC epitope from a protein encoded by the RD1 or RD2 regions of the
 CC Mycobacterium tuberculosis, M. bovis or M. africanum genome. The
 CC diagnostic reagents comprised within the kit are selected so that they
 CC are able to differentiate between M. bovis, M. tuberculosis or M.
 CC africanum-infected mammals and mammals vaccinated against M. bovis, M.
 CC tuberculosis or M. africanum. The polypeptide or its variant or fragment
 CC is useful as a medicament provided that the polypeptide is not a MPT-64
 CC polypeptide or a polypeptide encoded by the Rv1984c region of the M.
 CC bovis, M. tuberculosis or M. africanum genomes. The nucleic acid which
 CC encodes the polypeptide or the diagnostic reagents can be used as a
 CC vaccine, especially against Mycobacterium infections. This sequence
 CC corresponds to a protein used to derive the peptide epitopes used in the
 CC method of the invention.
 CC
 XX Sequence 99 AA;
 SQ

Query Match 100.0%; Score 479; DB 7; Length 99;
 Best Local Similarity 100.0%; Pred. No. 3.4e-46;
 Matches 98; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEKSHDPIADIGTGVSDNALHGVTAGSTALTSTVGLVPGADSVSAQAATAFTSSGIG 60
 |||||||
 DB 1 MEKSHDPIADIGTGVSDNALHGVTAGSTALTSTVGLVPGADSVSAQAATAFTSSGIG 60

QY 61 LLASNASADQDLHRAGBAVDVARTYSQIDGAGVFA 98
 |||||||
 DB 61 LLASNASADQDLHRAGBAVDVARTYSQIDGAGVFA 98

RESULT 5
 AAU70949 standard; protein; 102 AA.
 AAU70949
 XX
 AC AAU70949;
 XX
 DT 25-FEB-2002 (first entry)
 XX
 DE M. tuberculosis Rv0285 protein.
 XX
 KM Tuberculosis; tuberculostatic; antibacterial; vaccine; Rv0284; Rv0285;
 KM Rv0455C; Rv0569; Rv1195; Rv1386; Rv3477; Rv3878; Rv3879; MT3106.1;
 KM ORF13A; Rv0284c; Mycobacterium bovis; Mycobacterium africanum;
 KM BCG vaccine.
 XX
 OS Mycobacterium tuberculosis.
 XX
 PN W0200179274-A2.
 XX
 PD 25-OCT-2001.
 XX
 PF 19-APR-2001; 2001WO-DK000276.
 XX
 PR 19-APR-2000; 2000DK-00006666.
 PR 21-FEB-2001; 2001DK-00000283.
 XX
 PA (SPAT-) STATENS SERUM INST.
 XX
 PI Agger EM, Andersen P, Okkels LM, Wellingh K;
 XX
 DR WPI; 2002-061970/08.
 DR N-PSDB; AAS95786.
 XX
 PT New Mycobacterium tuberculosis antigens, useful for diagnosing
 PT tuberculosis, and as a vaccine for treating or preventing infections
 PT caused by species of tuberculosis complex.
 XX
 PS Claim 1; Page 73-74; 11pp; English.
 XX
 CC The invention relates to a substantially pure polypeptide comprising an
 CC amino acid sequence selected from Rv0284, Rv0285, Rv0455C, Rv0569,
 CC Rv1195, Rv1386, Rv3477, Rv3878, Rv3879c or MT3106.1 (also disclosed are
 CC ORF13A and Rv0284c), or their immunogenic portion, nucleic acids
 CC encoding them and an amino acid sequence analogue having at least 70%
 CC sequence identity to the polypeptide and is immunogenic. The protein is
 CC useful in preparing a pharmaceutical composition for diagnosing
 CC tuberculosis and in preparing a vaccine against tuberculosis caused by
 CC virulent mycobacteria. The vaccine or immunogenic/ pharmaceutical
 CC composition can be used prophylactically in a subject not infected with a
 CC virulent mycobacterium, or therapeutically in a subject already infected
 CC with a virulent mycobacterium. The protein is useful for preventing,
 CC treating and detecting infections caused by species of tuberculosis
 CC complex (M. tuberculosis, M. bovis, M. africanum). The nucleic acids may
 CC be used for effecting in vivo expression of the antigen, and in
 CC diagnostic assays for detecting the presence of pathogenic organisms in a
 CC sample. The vaccine is an improvement of the living BCG vaccine presently
 CC available, where one or more copies of the DNA sequence encoding one or
 CC more polypeptide has been incorporated into the genome of the
 CC microorganism to allow the microorganism to express and secrete the
 CC polypeptide. Incorporation of more than one copy of a nucleotide sequence
 CC enhances the immune response. The present sequence represents an M.
 CC tuberculosis protein of the invention
 XX
 SQ Sequence 102 AA;

useful for the preparation of an immunogenic composition, vaccine or pharmaceutical composition for a prophylactic or therapeutic vaccination against an infection caused by a virulent mycobacterium, where the immunogenetic composition is to be administered to a person which has previously been vaccinated with BCG. This sequence represents Mycobacterium tuberculosis Rv0285 polypeptide.

Sequence 102 AA:

	Query Match	21.4%	Score 102.5;	DB 9;	Length 102;
BcT	Local Similarity	29.2%;	Pred. No. 0.0015;		
Mtch	Matches	26; Conservative	17; Mismatches	41; Indels	5; Gaps

Dy
GTVSDNALHGVV-----AGSTALTSYTGVLPGADSVSAQAAFTPSGIQLAASNAS 68
| | :
GLAAASAVENTLRLRAAHASAPVITVAAPPADVPVSLOTANGFSGAQVEHAIVTTAE 68

Dz
69 QDOLRAGEAVOPDVARTYSOIDDGAAGVF 97
::|||::|::|::|::|::|::|::|::|:
69 VEHGRGVRGCVGESGASYLAGDAANAATRY 97

RESULT 7
ABE26444 ID ABE26444 standard; protein; 487 AA.
AC AEB26444;
AD 08-SEP-2005 (first entry)
DE M. tuberculosis fusion polypeptide, TB10.4-Rv0285-Ag85A.
EK immunogenicity; vaccine; pharmaceutical; fusion protein; Rv0285; TB10.4;
EL Ag85A; mycobacterium tuberculosis infection; infection; immunity;
EM tuberculostatic; mutain.
EN
EO Mycobacterium tuberculosis.
EP Synthetic.
EQ WO2005061534-A2.
ER
ES 07-JUL-2005.
ET 22-DEC-2004; 2004MO-DK00907.
EU 23-DEC-2003; 2003DK-00001942.
EV (STAT-) STATENS SERUM INST.
FX Des D, Andersen P, Agaard C,
FY WPI, 2005-479440/48.
FZ N-PSTD; ABB26443.
GX New immunogenic compositions, vaccines or pharmaceutical compositions having a fusion polypeptide with antigens from Bacillus Calmette Guérin (BCG) administered together with BCG.
GY Disclosure; SEQ ID NO 16; 85dp; English.

The invention relates to an immunogenic composition, vaccine or pharmaceutical composition comprising a fusion polypeptide having antigenic determinants derived from the fusion proteins described herein, wherein said antigenic determinants are selected from the group consisting of: (1) an immunogenic composition, vaccine or pharmaceutical composition comprising a fusion polypeptide which comprises two or more immunogenic polypeptides, or its analogue, selected from Ag85B, TB10.4, ORP2C, Rv0285, Rv0287 and Rv1036c, or CC comprises amino acid sequences encoding the fusion polypeptides selected from Ag85B-TB10.4, Ag85B-TB10.4-Ag85A, Ag85B-TB10.4-ORP2C, Ag85B-TB10.4-ORP2C-Ag85A, Ag85B-TB10.4-Rv1036, Ag85A-TB10.4, Ag85B-TB10.4-Ag85A-TB10.4, TB10.4-Rv0285-Ag85A, TB10.4-Rv1036-Ag85A, TB10.4-ORP2C-Ag85A, Ag85A-Rv0287, Rv0287-TB10.4, and their analogues; (2) a vaccine or

CC		pharmaceutical composition comprising a nucleic acid fragment, comprising
CC		a nucleotide sequence encoding a fusion polypeptide as mentioned; and (3)
CC		a method for immunizing and/or treating an animal, which has previously
CC		been immunized against tuberculosis caused by a virulent mycobacterium
CC		with BCG or having active or latent tuberculosis caused by a virulent
CC		mycobacterium, comprising administering to the animal any of the
CC		immunogenic compositions, vaccines or pharmaceutical compositions
CC		mentioned. The immunogenic composition, vaccine or pharmaceutical
CC		composition are useful for therapeutic use, prophylactic use, or for
CC		boosting immunity from prior BCG vaccination. The fusion polypeptide is
CC		useful for the preparation of an immunogenic composition, vaccine or
CC		pharmaceutical composition for a prophylactic or therapeutic
CC		vaccination against an infection caused by a virulent mycobacterium,
CC		where the immunogenic composition is to be administered to a person which
CC		has previously been vaccinated with BCG. This sequence represents a
CC		Mycobacterium tuberculosis fusion polypeptide.
XX		
XX	SQ	Sequence 487 AA;
OY		
Db		
OY	69	QDQLRAGEAVQDVARTYSQIDGAAGVF 97
Db	105	GTAASAIVKALTARLAAHNSAAPVITAVVPADPVSILQTAGSAGCSEHAVTALG 164
Db	165	VSELGRAGVGVESSGASYLAGMAAATY 193
RESULT 8		
ID	ABU36685	standard; protein; 639 AA.
AC	ABU36685;	
XX		
DT	19-JUN-2003	(first entry)
XX		
DE		Protein encoded by Prokaryotic essential gene #22212.
KW		Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX		
OS		Mycobacterium tuberculosis.
PN		WO200277183-A2.
XX		
PD	03-OCT-2002.	
XX		
PF	21-MAR-2002;	2002MO-US009107.
XX		
PR	21-MAR-2001;	2001US-00815242.
PR	06-SEP-2001;	2001US-00948993.
PR	25-OCT-2001;	2001US-0342923P.
PR	08-FEB-2002;	2002US-00072851.
PR	06-MAR-2002;	2002US-0362699P.
XX		
PA	(BLT-) ELITRA PHARM INC.	
XX		
P1	Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW,	
P1	Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;	
XX		
DR	WPI; 2003-029926/02.	
XX	N-PSDB; ACA40555.	
PT		New antisense nucleic acids, useful for identifying proteins or screening
PT		for homologous nucleic acids required for cellular proliferation to
PT		isolate candidate molecules for rational drug discovery programs.
XS		
CS	Claim 25; SEQ ID NO 64609; 1766pp; English.	
CC		The invention relates to an isolated nucleic acid comprising any one of

the 623 antisense sequences given in the specification where expression of the nucleic acid inhibits proliferation of a cell. Also included are: (1) a vector comprising a promoter operably linked to the nucleic acid encoding a polypeptide whose expression is inhibited by the antisense nucleic acid; (2) a host cell containing the vector; (3) an isolated polypeptide or its fragment whose expression is inhibited by the antisense nucleic acid; (4) an antibody capable of specifically binding the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular proliferation or the activity of a gene in an operon required for proliferation; (7) identifying a compound that influences the activity of the gene product or that has an activity against a biological pathway required for proliferation, or that inhibits cellular proliferation; (8) identifying a gene required for cellular proliferation or the biological pathway in which a proliferation-regulated gene or its gene product lies or a gene on which the best compound that inhibits proliferation of an organism acts; (9) manufacturing an antibiotic; (10) profiling a compound's activity; (11) a culture comprising strains in which the gene product is overexpressed or underexpressed; (12) determining the extent to which each of the strains is present in a culture or collection of strains; or (13) identifying the target of a compound that inhibits the proliferation of an organism. The antisense nucleic acids are useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational drug discovery programs, or for screening homologous nucleic acids required for proliferation in cells other than *S. aureus*, *S. typhimurium*, *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of the target prokaryotic essential genes. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at [ftp.wipo.int/pub/published_pct_sequences](http://wipo.int/pub/published_pct_sequences)

```
Query Match      19.3%; Score 92.5; DB 6; Length 639;
Beet Local Similarity 27.1%; Pred. No.0.23;
Matches       32; Conservative    15; Mismatches     32; Indels   39; Gaps    4
```

Dy | | : | | : | | : | | : | | : | | : | | : | | : |
 ADIGTQVSDNALHGVTAGSTALTTSYTGAVPAPGADDEVSAOATAFTSEGI-----58
Db 21 ATISASIS-----RAVAEALVPITALLIPGADVSNAILAFATHTGOAYQLSHAWA 73

Dy 59 -----IOLLASNASH-----ODLHHRAGEANVDVARYSQ-----IDDGAAGV 96
 -:::-:::-:::-:::-:::-:::-:::-:::-:::-:
Db 74 FHEGFVLMSGAQAQXASAEEANSPLDIYGQTALDAINSPQTLGRFLONGANGV 131

RESULT 9
AEB91455
ID AEB91455 standard; protein; 639 AA.
XX AC AEB91455;
XX DT 20-OCT-2005 (first entry)
XX DB Microbial pathogen adhesion protein sequence, SEQ ID NO:165.
XX KX algorithm; adhesin; pharmaceutical; vaccine; drug screening;
KW bordetella pertussis infection; antibacterial; pneumonia;
KV antiinflammatory; respiratory-gen.; gastric ulcer; antilucer;
KM gastrointestinal-gen.; urinary tract infection; antimicrobial; uropathic.

OS Mycobacterium tuberculosis.
PN MO2005076010-A2.
PD 18-AUG-2005.
PF 07-FEB-2005; 2005WO-IN000037.
PR 06-FEB-2004; 2004IN-DRO00173.
RR 20-JUL-2004; 2004US-05589227F.
XX (COUNCIL SCI & IND RES SOUTH AFRICA.)

XX Sachdeva G, Kumar K, Jain P, Brahmachari SK, Ramachandran S;
PI WPI; 2005-597835/61.
DR
XX
XX Computational method for identifying adhesin and adhesin like molecules,
PT computes computing sequence-based attributes of protein sequences using
PT neural network software and training an artificial neural network.
XX
PS Claim 16; SEQ ID NO 165; 402pp; English.
XX
CC The present invention relates to a computational method (M1) for
CC identifying adhesin and adhesin-like proteins, by computing the sequence-
CC based attributes of protein sequences using five attribute modules of a
CC neural network software, training an artificial neural network (ANN) for
CC each of the computed five attributes, and identifying the adhesin and
CC adhesin-like proteins having probability of being an adhesin (pad) as
CC equal or greater than 0.51. Also claimed is a set of 274 annotated genes
CC encoding adhesin and adhesin-like proteins, having 274 fully defined 162-
CC 1151 base pair (SEQ ID NO: 385-658) sequences; a set of 105 hypothetical
CC genes encoding adhesin and adhesin-like proteins, having 105 fully
CC defined 306-15876 base pairs (SEQ ID NO: 659-763); a set of 279 annotated
CC adhesin and adhesin-like proteins, having 279 fully defined 53-3716 base
CC pair (SEQ ID NO: 1-279); a set of 105 hypothetical adhesin and adhesin-
CC like proteins, having 105 fully defined 106-5291 base pair (SEQ ID NO:
CC 280-384) sequences; and a fully connected multilayer feed forward ANN (1)
CC based on (M1). (M1) is useful for identifying adhesin and adhesin-like
CC proteins, of therapeutic potential, and identifying and short-listing
CC proteins for further testing in development of new vaccine formulations
CC to eliminate diseases caused by various pathogenic organisms. (M1) is
CC useful for identifying putative adhesins that are important in drug
CC discovery and preventing therapeutics for whooping cough, pneumonia,
CC gastric ulcer and urinary tract infections. (M1) identifies adhesins from
CC distantly related organisms, and from bacteria belonging to a wide
CC phylogenetic spectrum. (M1) is capable of predicting adhesive nature of
CC unique proteins. The present sequence is a microbial pathogen adhesin
CC protein sequence.
XX
XX Sequence 639 AA;
XX
Query Match 19.3%; Score 92.5; DB 9; Length 639;
Best Local Similarity 27.1%; Pred. No. 0.23; Indels 39; Gaps 4;
Matches 32; Conservative 15; Mismatches 32; Indels 39; Gaps 4;
QY 11 ADIGTQVSDNALHGVATGTAATVTGVPAGADEVSAQAATFTSGE----- 58
DB 21 ATIGSAIS-----RANNAENLVPTTLIPAGADVDVAALAAAPATGGAQVGLSANA 73
QY 59 -----IQLLASNNSA-----ODQLHRAGEAVQDVARTYSQ-----IDDGACV 96
DB 74 FHQGFVQLMGAAGAAQVASAEANSSPLQIVGQRTALDAINSPVQTLTGRLPIGNGANGV 131

RESULT 10
ABU36945
ID ABU36945 standard; protein; 588 AA.
XX
XX ABU36945;
AC
XX 19-JUN-2003 (first entry)
DT
XX
DE Protein encoded by Prokaryotic essential gene #22472.
XX
XX Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX
XX Mycobacterium tuberculosis.
OS
XX
XX WO200277183-A2.
PN
XX
PD 03-OCT-2002.
PP
XX 21-MAR-2002; 2002WO-US009107.
XX

PR 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-03626599P.
XX
XX (ELITR) ELITRA PHARM INC.
XX
XX Wang L, Zamudio C, Malone C, Haseelbeck R, Ohlsen KL, Zykkind JW,
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX
XX WPI; 2003-029926/02.
DR
XX
XX N-PSDB; AQA40815.
XX
XX New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
XX
XX Claim 25; SEQ ID NO 64869; 1766pp; English.
XX
XX The invention relates to an isolated nucleic acid comprising any one of
CC the 623 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 588 AA;
XX
Query Match 18.8%; Score 90; DB 6; Length 588;
Best Local Similarity 33.0%; Pred. No. 0.39; Indels 24; Gaps 4;
Matches 30; Conservative 13; Mismatches 24; Indels 24; Gaps 4;
QY 11 ADIGTQVSDNALHGVATGTAATVTGVPAGADEVSAQAATFTSGE--IQLLASNNSA 68
DB 19 ANIGSALS-----RANNAENLVPTTLIPAGADVDVAALAAAPATGGAQVGLSANA 71
QY 69 -ODQLHRAGEAVQDVARTYSQIDGACVFA 98
DB 72 LHDQ-----FVQALTGAGGSA 88

RESULT 11
ADM40790
ID ADM40790 standard; protein; 588 AA.
XX
XX ADM40790;
AC

[illegible]

XX	MO200277183-A2.	
XX		
PN	03-OCT-2002.	
PD		
XX		
XX	21-MAR-2002; 2002MO-US009107.	
PF		
XX		
PR	21-MAR-2001; 2001US-00815242.	
PR	06-SEP-2001; 2001US-00948993.	
PR	25-OCT-2001; 2001US-0342923P.	
PR	08-FEB-2002; 2002US-00072851.	
PR	06-MAR-2002; 2002US-0362699P.	
XX		
XX	(ELITR) ELITRA PHARM INC.	
PA		
PI	Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;	
PI	Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;	
XX		
XX	WPI. 2003-029926/02.	
DR	N-PSDB; AC40410.	
PT		
PT	New antisense nucleic acids, useful for identifying proteins or screening	
PT	for homologous nucleic acids required for cellular proliferation to	
PT	isolate candidate molecules for rational drug discovery programs.	
XX		
XX	Claim 25; SEQ ID NO 64464; 1766pp; English.	
PS		
XX		
CC	The invention relates to an isolated nucleic acid comprising any one of	
CC	the 6213 antisense sequences given in the specification where expression	
CC	of the nucleic acid inhibits proliferation of a cell. Also included are:	
CC	(1) a vector comprising a promoter operably linked to the nucleic acid	
CC	encoding a polypeptide whose expression is inhibited by the antisense	
CC	nucleic acid; (2) a host cell containing the vector; (3) an isolated	
CC	polypeptide or its fragment whose expression is inhibited by the	
CC	antisense nucleic acid; (4) an antibody capable of specifically binding	
CC	the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular	
CC	proliferation or the activity of a gene in an operon required for	
CC	proliferation; (7) identifying a compound that influences the activity of	
CC	the gene product or that has an activity against a biological pathway	
CC	required for proliferation, or that inhibits cellular proliferation; (8)	
CC	identifying a gene required for cellular proliferation or the biological	
CC	pathway in which a proliferation-required gene or its gene product lies	
CC	or a gene on which the test compound that inhibits proliferation of an	
CC	organism acts; (9) manufacturing an antibiotic; (10) profiling a	
CC	compound's activity; (11) a culture comprising strains in which the gene	
CC	product is overexpressed or underexpressed; (12) determining the extent	
CC	to which each of the strains is present in a culture or collection of	
CC	strains; or (13) identifying the target of a compound that inhibits the	
CC	proliferation of an organism. The antisense nucleic acids are useful for	
CC	identifying proteins or screening for homologous nucleic acids required	
CC	for cellular proliferation to isolate candidate molecules for rational	
CC	drug discovery programs, or for screening homologous nucleic acids	
CC	required for proliferation in cells other than <i>S. aureus</i> , <i>S. typhimurium</i> ,	
CC	<i>K. pneumoniae</i> or <i>P. aeruginosa</i> . The present sequence is encoded by one of	
CC	the target prokaryotic essential genes. Note: The sequence data for this	
CC	parent did not form part of the primed specification, but was obtained	
CC	in electronic format directly from WPO at	
CC	ftp.wipo.int/pub/published_pct_sequences	
XX		
XX	Sequence 606 AA;	
SQ		
Query Match	18.3%; Score 87.5; DB 6; Length 606;	
Best Local Similarity	32.9%; Pred. NO. 0.79;	
Matches	27; Conservative 15; Mismatches 29; Indels 11; Gaps 3	
0Y	11 ADIGTQVSDNALHGVTAGSTALTSVTGLVPGADDEVSAQAATAFT--SEGIQLASNSA 68	
DB	19 AIGISTLS-----AANAAALAPFTGVLTAAGDADEVSAVASLFSGHQAQVOTLCTGAAA 71	
0Y	69 ODQLHRAGEAVQDVARTYSQID 90	
DB	72 FHE--RFIQLSTAGAGASAE 91	

01	RESULT 13
02	AEB91453
03	ID AEB91453 standard; protein; 606 AA.
04	XX
05	AC AEB91453;
06	XX
07	XX 20-OCT-2005 (first entry)
08	XX
09	DE Microbial pathogen adhesin protein sequence, SEQ ID NO:163.
10	XX
11	KW algorithm; adhesin; pharmaceutical; vaccine; drug screening;
12	KM borreliella pertussis infection; antibacterial; pneumonia;
13	KM antiinflammatory; respiratory-gen.; gastric ulcer; antidiacer;
14	XX gastrointestinal-gen.; urinary tract infection; antimicrobial; uropathic.
15	XX
16	OS Mycobacterium tuberculosis.
17	XX
18	PN WO2005076010-A2.
19	XX
20	PD 18-AUG-2005.
21	XX
22	PF 07-FEB-2005; 2005WO-IN0000037.
23	XX
24	PR 06-FEB-2004; 2004IN-DE0000173.
25	XX
26	PR 20-JUL-2004; 2004US-0589227P.
27	XX
28	PA (COUL) COUNCIL SCI & IND RES SOUTH AFRICA.
29	XX
30	P1 Sachdeva G, Kumar K, Jain P, Brahmachari SK, Ramchandran S;
31	XX
32	DR WPI, 2005-597835/61.
33	XX
34	PT Computational method for identifying adhesin and adhesin like molecules,
35	PT comprises computing sequence-based attributes of protein sequences using
36	PT neural network software and training an artificial neural network.
37	XX
38	PS Claim 16; SEQ ID NO 163; 402pp; English.
39	XX
40	CC The present invention relates to a computational method (M1) for
41	CC identifying adhesin and adhesin-like proteins, by computing the sequence-
42	CC based attributes of protein sequences using five attribute modules of a
43	CC neural network software, training an artificial neural network (ANN) for
44	CC each of the computed five attributes, and identifying the adhesin and
45	CC adhesin-like proteins having probability of being an adhesin (Pad) as
46	CC equal or greater than 0.51. Also claimed is a set of 274 annotated genes
47	CC encoding adhesin and adhesin-like proteins, having 274 fully defined 162-
48	CC 1131 base pair (SEQ ID NO: 385-658) sequences; a set of 105 hypothetical
49	CC genes encoding adhesin and adhesin-like proteins, having 105 fully
50	CC defined 306-15876 base pairs (SEQ ID NO: 659-763); a set of 279 annotated
51	CC adhesin and adhesin-like proteins, having 279 fully defined 53-7716 base
52	CC pair (SEQ ID NO: 1-279); a set of 105 hypothetical adhesin and adhesin-
53	CC like proteins, having 105 fully defined 106-5291 base pair (SEQ ID NO:
54	CC 280-384) sequences; and a fully connected multilayer feed forward ANN (I)
55	CC based on (M1). (M1) is useful for identifying adhesin and adhesin-like
56	CC proteins, of therapeutic potential, and identifying and short-listing
57	CC proteins for further testing in development of new vaccine formulations
58	CC to eliminate diseases caused by various pathogenic organisms. (M1) is
59	CC useful for identifying putative adhesins that are important in drug
60	CC discovery and preventing therapeutics for whooping cough, pneumonia,
61	CC gastric ulcer and urinary tract infections. (M1) identifies adhesins from
62	CC distantly related organisms, and from bacteria belonging to a wide
63	CC phylogenetic spectrum. (M1) is capable of predicting adhesive nature of
64	CC unique proteins. The present sequence is a microbial pathogen adhesin
65	CC protein sequence.
66	XX
67	XX Sequence 606 AA;
68	XX
69	Query Match 18.3%; Score 87.5; DB 9; Length 606;
70	Best Local Similarity 32.9%; Pred. No. 0.79;
71	Matches 27; Conservative 15; Mismatches 29; Indels 11; Gaps 3;
72	11 ADIGQVSDNALHGVTAAGSTALTSTVTGLVPAGADVSAQAATATFT--SEGIQILASNSA 68

Db 19 AQTSTLS-----AANAAALATTTGTLAAGDVEAANAVALSLSGHQAQVTLGTQAAA 71
 Oy 69 QDQLRAGEAVQDVARTYSQID 90
 Db 72 FHE--RPIQALSTAAAGVSAE 91
 RESULT 14
 AAU70953
 ID AAU70953 standard; protein, 102 AA.
 AC AAU70953;
 XX
 XX 25-FEB-2002 (first entry)
 DE M. tuberculosis Rv1386 protein.
 XX
 XX Tuberculosis; Tuberculoestic; antibacterial; vaccine; Rv0284; Rv0285;
 KW Rv0455c; Rv1386; Rv1386; Rv3477; Rv3878; Rv3879; MT3106..1;
 KM ORF13A; Rv0284c; Mycobacterium bovis; Mycobacterium africanum;
 KM BCG vaccine.
 OS Mycobacterium tuberculosis.
 XX
 XX WO200179274-A2.
 PN 25-OCT-2001.
 PD
 XX
 XX 19-APR-2001; 2001WO-DK000276.
 PF
 XX 19-APR-2000; 2000DK-00000666.
 PR 21-FEB-2001; 2001DK-00000283.
 XX
 PA (STAT-) STATENS SERUM INST.
 PI Agger EM, Andersen P, Okkeis LMW, Weidlingh K;
 XX
 DR MPI; 2002-061970/08.
 XX
 DR N-PSDB; AAS95790.
 PT
 PT New Mycobacterium tuberculosis antigens, useful for diagnosing
 PT tuberculosis, and as a vaccine for treating or preventing infections
 PT caused by species of tuberculosis complex.
 XX
 PT Claim 1; Page 77; 11pp; English.
 PS
 CC The invention relates to a substantially pure polypeptide comprising an
 CC amino acid sequence selected from Rv0284, Rv0285, Rv0455c, Rv0569,
 CC Rv1195, Rv1386, Rv3477, Rv3878, Rv3879 or MT3106.1 (also disclosed are
 CC ORF13A and Rv0284c), or their immunogenic portion, nucleic acids
 CC encoding them and an amino acid sequence analogue having at least 70%
 CC sequence identity to the polypeptide and is immunogenic. The protein is
 CC useful in preparing a pharmaceutical composition for diagnosing
 CC tuberculosis and in preparing a vaccine against tuberculosis caused by
 CC virulent mycobacteria. The vaccine or immunogenic/ pharmaceutical
 CC composition can be used prophylactically in a subject not infected with a
 CC virulent mycobacterium, or therapeutically in a subject already infected
 CC with a virulent mycobacterium. The protein is useful for preventing,
 CC treating and detecting infections caused by species of tuberculosis
 CC complex (M. tuberculosis, M. bovis, M. africanum). The nucleic acids may
 CC be used for effecting in vivo expression of the antigen, and in
 CC diagnostic assays for detecting the presence of pathogenic organisms in a
 CC sample. The vaccine is an improvement of the living BCG vaccine presently
 CC available, where one or more copies of the DNA sequence encoding one or
 CC more polypeptide has been incorporated into the genome of the
 CC microorganism to allow the microorganism to express and secrete the
 CC polypeptide. Incorporation of more than one copy of a nucleotide sequence
 CC enhances the immune response. The present sequence represents an M.
 CC tuberculosis protein of the invention
 XX
 XX Sequence 102 AA;
 SQ

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: April 14, 2006, 17:19:08 ; Search time 5.7465 Seconds
(without alignments)
1640.866 Million cell updates/sec

Title: US-10-620-246-88

Perfect score: 479
Sequence: 1 MEKMSHDPPIADIGTQVSDN.....VQDVARTYSQIDGAGVFA 98

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : PIR_80:*
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	479	100.0	99	2 F70802	probable PB protei
2	231	48.2	111	2 C70789	probable PB protei
3	102.5	21.4	102	2 C70836	probable PB protei
4	92.5	19.3	639	2 D70931	hypothetical glyci
5	90	18.8	588	2 F70971	hypothetical glyci
6	87.5	18.3	606	2 H70816	hypothetical glyci
7	87	18.2	102	2 H70898	probable PB protei
8	86.5	18.1	783	2 E70824	hypothetical glyci
9	86.5	18.1	1489	2 D70807	hypothetical glyci
10	86.5	18.1	1901	2 F70806	hypothetical glyci
11	85.5	17.8	543	2 F70726	hypothetical glyci
12	85.5	17.8	801	2 F70824	hypothetical glyci
13	85	17.7	487	2 E70983	hypothetical glyci
14	83	17.3	576	2 A70900	hypothetical glyci
15	82	17.1	562	2 B70953	hypothetical glyci
16	81	16.9	397	2 A72870	HLVD family secret
17	81	16.9	403	2 H97646	probable membrane
18	81	16.9	1011	2 F70620	hypothetical glyci
19	80.5	16.8	853	2 A70896	hypothetical glyci
20	80.5	16.8	914	2 H70987	hypothetical glyci
21	80.5	16.8	1660	2 A70889	hypothetical glyci
22	80	16.7	588	2 B70618	probable PB protei
23	80	16.7	767	2 E70895	hypothetical glyci
24	80	16.7	923	2 E70820	hypothetical glyci
25	79.5	16.6	370	2 G70872	hypothetical glyci
26	79.5	16.6	882	2 B70812	hypothetical glyci
27	78.5	16.4	449	2 F84241	hypothetical prote
28	78.5	16.4	957	2 D70835	hypothetical glyci
29	78	16.3	102	2 C87226	PB-family protein

30	78	16.3	545	2 T46811	halobacterial tran
31	78	16.3	545	2 E84327	Hcr7 transducer (l
32	77.5	16.2	99	2 F70929	probable PB protei
33	77.5	16.2	603	2 A70770	hypothetical glyci
34	77	16.1	463	2 B70893	hypothetical glyci
35	76.5	16.0	439	2 D70954	hypothetical glyci
36	76	15.9	94	2 C70564	probable PB protei
37	76	15.9	544	2 T44938	transducer protein
38	75.5	15.9	1467	2 A75564	conserved hypotet
39	75.5	15.8	100	2 S72816	hypothetical prote
40	75.5	15.8	558	2 E70756	hypothetical glyci
41	75.5	15.8	615	2 H70589	hypothetical glyci
42	75.5	15.8	1306	2 A70934	hypothetical glyci
43	75	15.7	99	2 G70560	probable PB protei
44	75	15.7	102	2 B86976	PB-family protein
45	75	15.7	594	2 G70545	hypothetical glyci

ALIGNMENTS

RESULT 1
F70802
probable PB protein - Mycobacterium tuberculosis (strain H37RV)
C/Species: Mycobacterium tuberculosis
C/Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999
C/Accession: F70802
R/Colo, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, J.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A/Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A/Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A/Reference number: A70500; MUID:98295987; PMID:9634230
A/Accession: F70802
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-99 <COL>
A/Cross-references: UNIPARC:UPI00000C1571, GB:AL022120, GB:AL123456, NID:93261558, PIDN
A/Experimental source: strain H37RV
A/Genetic8:
A/Gene: PB

Query Match 100.0%; Score 479; DB 2; Length 99;
Best Local Similarity 100.0%; Pred. No. 3e-37;
Matches 99; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEKMSHDPPIADIGTQVSDNALHGVTAAGTALTSVTGLVPAGADVEVSAQAATAFTSGIQ 60
DB 1 MEKMSHDPPIADIGTQVSDNALHGVTAAGTALTSVTGLVPAGADVEVSAQAATAFTSGIQ 60
QY 61 LLASMSAODOLHRAEAVQDVARTYSQIDGAGVFA 98
DB 61 LLASMSAODOLHRAEAVQDVARTYSQIDGAGVFA 98

RESULT 2
C70789
probable PB protein - Mycobacterium tuberculosis (strain H37RV)
C/Species: Mycobacterium tuberculosis
C/Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C/Accession: C70789
R/Colo, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, J.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A/Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A/Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A/Reference number: A70500; MUID:98295987; PMID:9634230
A/Accession: C70789
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-111 <COL>

A/Cross-references: UNIPROT:O69713; UNIPARC:UPI000016537F; GB:AL022121; GB:AL123456; NID:
 A/Experimental source: strain H37Rv
 C/Genetics:
 A/Gene: PE

Query Match 48.2%; Score 231; DB 2; Length 111;
 Best Local Similarity 52.1%; Pred. No. 1.9e-14;
 Matches 49; Conservative 15; Mismatches 30; Indels 0; Gaps 0;
 Oy 1 MEMKSHDPIADIGTOVSDNALHGVTGTLTSTGTVPGADSVSAQAATFTSEGIQ 60
 Db 1 MGSNFPDPAVADIGSVVNNAFQGLQAGAVANWSLSLIPAGAEVSAKAVTATTATG 60
 Oy 61 LLASNSAQDOLHRAGEAVQDVARTYSQIDGAA 94
 Db 61 LLALNQAQOEELRKAGVFTAIAMYSADADYRAA 94

RESULT 3

C70836
 C/Species: Mycobacterium tuberculosis (strain H37Rv)
 C/Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004

C/Accession: C70836
 R/Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
 Nature 393, 537-544, 1998
 A/Authors: Sgaree, R.; Sultston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
 A/Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
 A/Reference number: A70500; PMID:98295987; PMID:9634220
 A/Accession: C70836
 A/Status: preliminary; nucleic acid sequence not shown; translation not shown
 A/Molecule type: DNA
 A/Residues: 1-102 <COL>
 A/Cross-references: UNIPROT:O53690; UNIPARC:UPI000043HB2; GB:AL021930; GB:AL123456; NID:
 A/Experimental source: strain H37Rv
 C/Genetics:
 A/Gene: PE

Query Match 21.4%; Score 102.5; DB 2; Length 102;
 Best Local Similarity 29.2%; Pred. No. 0.011;
 Matches 26; Conservative 17; Mismatches 41; Indels 5; Gaps 1;
 Oy 14 GTGVSDNALHGVT-----AGSTALTSTGTVPGADSVSAQAATFTSEGIQLASNSA 68
 Db 9 GLAASAAVAETRLRLAAHSAAPVITAVPPAADPVSLQTAAAGFAGVHNVAVTREG 68
 Oy 69 QDOLHRAGEAVQDVARTYSQIDGAAQVF 97
 Db 69 VEEIGRAGVGIVGEGASVYLAGDAATAATY 97

RESULT 4

D70931
 C/Species: Mycobacterium tuberculosis
 C/Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 31-Mar-2003

C/Accession: D70931
 R/Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
 Nature 393, 537-544, 1998
 A/Authors: Sgaree, R.; Sultston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
 A/Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
 A/Reference number: A70500; PMID:98295987; PMID:9634220
 A/Accession: D70931
 A/Status: preliminary; nucleic acid sequence not shown; translation not shown
 A/Molecule type: DNA
 A/Residues: 1-639 <COL>
 A/Cross-references: UNIPARC:UPI0000003B03; GB:AL022021; GB:AL123456; NID:G3250699; PIDN:
 A/Experimental source: strain H37Rv
 C/Genetics:

A/Gene: RV1803c
 C/Superfamily: uncharacterized glycine-rich protein, PE motif containing

Query Match 19.3%; Score 92.5; DB 2; Length 639;
 Best Local Similarity 27.1%; Pred. No. 0.74;
 Matches 32; Conservative 15; Mismatches 32; Indels 39; Gaps 4;
 Oy 11 ADIGTOVSDNALHGVTGTLTSTGTVPGADSVSAQAATFTSEGIQ 58
 Db 21 ATIGSAIS-----PAMNALTPIPALIPAGADVDVSAIIMLFAITGQAVQELSAHVA 73
 Oy 59 -----IQLASNSA-----QDOLHRAGEAVQDVARTYSQ-----IDGAAQV 96
 Db 74 FHEQFQLMSAGAAQVSAEANSSPQIVGQALPDINSPPVTLTGRPLIGANGV 131

RESULT 5

F70971
 C/Species: Mycobacterium tuberculosis
 C/Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 31-Mar-2003

C/Accession: F70971
 R/Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
 Nature 393, 537-544, 1998
 A/Authors: Sgaree, R.; Sultston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
 A/Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
 A/Reference number: A70500; PMID:98295987; PMID:9634220
 A/Accession: F70971
 A/Status: preliminary; nucleic acid sequence not shown; translation not shown
 A/Molecule type: DNA
 A/Residues: 1-588 <COL>
 A/Cross-references: UNIPARC:UPI000004F71; GB:AL009198; GB:AL123456; NID:G3242262; PIDN:
 A/Experimental source: strain H37Rv
 C/Genetics:
 A/Gene: RV3367
 C/Superfamily: uncharacterized glycine-rich protein, PE motif containing

Query Match 18.8%; Score 90; DB 2; Length 588;
 Best Local Similarity 33.0%; Pred. No. 1.1;
 Matches 30; Conservative 13; Mismatches 24; Indels 24; Gaps 4;
 Oy 11 ADIGTOVSDNALHGVTGTLTSTGTVPGADSVSAQAATFTSEGI--QLASNSA 68
 Db 19 ANIGSALS-----AANAAAGTTGLLAAGDEVSALASIFSGHAYSQCVAAQATA 71
 Oy 69 -QDOLHRAGEAVQDVARTYSQIDGAAQVRA 98
 Db 72 LHDQ-----FVQALITGAGGSYA 88

RESULT 6

H70816
 C/Species: Mycobacterium tuberculosis
 C/Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 31-Mar-2003

C/Accession: H70816
 R/Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
 Nature 393, 537-544, 1998
 A/Authors: Sgaree, R.; Sultston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
 A/Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
 A/Reference number: A70500; PMID:98295987; PMID:9634220
 A/Accession: H70816
 A/Status: preliminary; nucleic acid sequence not shown; translation not shown
 A/Molecule type: DNA
 A/Residues: 1-606 <COL>
 A/Cross-references: UNIPARC:UPI0000003AFD; GB:AL022004; GB:AL123456; NID:G3261550; PIDN:
 A/Experimental source: strain H37Rv
 C/Genetics:
 A/Gene: RV0872c

C:Superfamily: uncharacterized glycine-rich protein, PB motif containing

Query Match 18.3%; Score 87.5; DB 2; Length 606;
Best Local Similarity 32.9%; Pred. No. 2;
Matches 27; Conservative 15; Mismatches 29; Indels 11; Gaps 3;

QY 11 ADICTQVSDNALHGVTAGSTALTSTVGLVPAGADEVSAQAATAFT-SEGIQLLASNSAQ 68
DB 19 AIGSTLS-----AANAALAPFTGVTLAGADEVSAVAALFSGHAQAVQTLCGTQAAA 71

QY 69 QDQHRAGEAVQDVARTYSQID 90
DB 72 FHR--RFLQALSTAGAGASAR 91

RESULT 7

H70898 Probable PB protein - Mycobacterium tuberculosis (strain H37RV)

C:Species: Mycobacterium tuberculosis

C>Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004

C:Accession: H70898

R:Coile, S.T.; Broesch, R.; Parhill, J.; Garner, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.

Nature 393, 537-544, 1998

A:Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.

A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome

A:Reference number: A70500; PMID:98295987; PMID:9634230

A:Accession: H70898

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-102 <COL>

A:Cross-references: UNIPROT:P71656; UNIPARC:UPI000011D867; GB:Z80108; GB:AL123456; NID:9

A:Experimental source: strain H37RV

C:Genetics:

A:Gene: PB

Query Match 18.2%; Score 87; DB 2; Length 102;
Best Local Similarity 26.6%; Pred. No. 0.29;
Matches 25; Conservative 18; Mismatches 39; Indels 12; Gaps 1;

QY 16 QVSDNALHGVTAGSTALTSTVGLVPAGADEVSAQAATAFTSGIQLLA 63
DB 4 RVPEESLAGSAALAEAVTARLAHAAPFAIAVPIPGSDSVSCNIVERSVAGSGIVA 63

QY 64 SNASADQHLRAGEAVQDVARTYSQIDGAGVF 97
DB 64 MAAGVERLGRSGVAGSASVARDALAAASY 97

RESULT 8

H70824 Hypothetical glycine-rich protein RV0746 - Mycobacterium tuberculosis (strain H37RV)

C:Species: Mycobacterium tuberculosis

C>Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004

C:Accession: E70824

R:Coile, S.T.; Broesch, R.; Parhill, J.; Garner, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.

Nature 393, 537-544, 1998

A:Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.

A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome

A:Reference number: A70500; PMID:98295987; PMID:9634230

A:Accession: E70824

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-783 <COL>

A:Cross-references: UNIPROT:O51809; UNIPARC:UPI00000D4FCA; GB:AL021958; GB:AL123456; NID

C:Experimental source: strain H37RV

C:Genetics:

A:Gene: RV0746

C:Superfamily: elastin related uncharacterized glycine-rich protein, PB motif containing

Query Match 18.1%; Score 86.5; DB 2; Length 783;
Best Local Similarity 30.3%; Pred. No. 3.3;
Matches 27; Conservative 14; Mismatches 45; Indels 3; Gaps 1;

QY 13 IGTVSDNALHGVTAGSTALTSTVGLVPAGADEVSAQAATAFTSGIQLLASNSAQ 69
DB 11 LGSATDIALAGSVLAGAADAAMAAATTTGIVAAQDEVSAAIYALFSAHGRAYQVVASQA 70

QY 70 DQHRAGEAVQDVARTYSQIDGAGVFA 98
DB 71 AVHAQFVHALSAGAGATASAAAGAVLA 99

RESULT 9

D70807 Hypothetical glycine-rich protein RV3514 - Mycobacterium tuberculosis (strain H37RV)

C:Species: Mycobacterium tuberculosis

C>Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000

C:Accession: D70807

R:Coile, S.T.; Broesch, R.; Parhill, J.; Garner, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.

Nature 393, 537-544, 1998

A:Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.

A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome

A:Reference number: A70500; PMID:98295987; PMID:9634230

A:Accession: D70807

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-1489 <COL>

A:Cross-references: UNIPARC:UPI00000D3AEB; GB:AL022022; GB:AL123456; NID:93261554; PIDN

A:Experimental source: strain H37RV

C:Genetics:

A:Gene: RV3514

C:Superfamily: collagen alpha 1(IV) chain

Query Match 18.1%; Score 86.5; DB 2; Length 1489;
Best Local Similarity 28.9%; Pred. No. 7;
Matches 24; Conservative 16; Mismatches 34; Indels 7; Gaps 1;

QY 12 DIGTVSDNALHGVTAGSTALTSTVGLVPAGADEVSAQAATAFTSGIQLLASNSAQDQ 71
DB 20 NLGSSIS-----AANAASATTVTLAAGADEVSAARIALFGFGLRYQALSAQVAAV 72

QY 72 LHRAGEAVQDVARTYSQIDGGA 94
DB 73 HQRVQALSTGAGVASAEAAA 95

RESULT 10

F70806 Hypothetical glycine-rich protein RV3508 - Mycobacterium tuberculosis (strain H37RV)

C:Species: Mycobacterium tuberculosis

C>Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004

C:Accession: F70806

R:Coile, S.T.; Broesch, R.; Parhill, J.; Garner, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.

Nature 393, 537-544, 1998

A:Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.

A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome

A:Reference number: A70500; PMID:98295987; PMID:9634230

A:Accession: F70806

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-1901 <COL>

A:Cross-references: UNIPROT:O53553; UNIPARC:UPI000013C2A6; GB:AL022022; GB:AL123456; NI

A:Experimental source: strain H37RV

C:Genetics:

A:Gene: RV3508

C:Superfamily: collagen alpha 1(IV) chain

Query Match 18.1%; Score 86.5; DB 2; Length 1901;

QY 69 QDQLHRAGEAVQDVARTYSQIDGAGVFA 98
 Db 72 FHQ-----QIVQTLTST-----AGAXA 88

RESULT 15

B70953
 hypothetical glycine-rich protein RV1243c - Mycobacterium tuberculosis (strain H37RV)
 C/Species: Mycobacterium tuberculosis
 C/Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
 C/Accession: B70953
 R/Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S. Nature 393, 537-544, 1998
 A/Authors: Sgaree, R., Sulston, J.E., Taylor, K., Whitehead, S., Barrell, B.G.
 A/Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
 A/Reference number: A70500; MUID:98295987; PMID:9634230
 A/Accession: B70953
 A/Status: preliminary; nucleic acid sequence not shown; translation not shown
 A/Molecule type: DNA
 A/Residues: 1562 <COL>
 A/Cross-references: UNIPROT:O50458; UNIPARC:UPI0000165283; GB:AL021006; GB:AL123456; NIT
 A/Experimental source: strain H37RV
 C/Genetics:
 A/Gene: RV1243c
 C/superfamily: uncharacterized glycine-rich protein, PB motif containing

Query Match 17.1%; Score 82; DB 2; Length 562;
 Best Local Similarity 31.6%; Pred. No. 5.9;
 Matches 37; Conservative 13; Mismatches 29; Indels 38; Gaps 6;
 QY 10 AAD---IGTQVSDNALHGVTAAGTALTSVTGLVPAGADVSAQAATAFT--SEGIQLLAS 64
 Db 15 AADLEGISALA-----AANRAAEAPTTGLLAAGADVSAATASLFSGNAQAYQALSA 67
 QY 65 NASA-QDQLHRA-----GEAVQDVARTYSQ-----IDGGAAG 95
 Db 68 QAAAFHQFVRAALLSAGSYAAAEANAASPMQAVLDVYVNGPTQLLGRPLIGDANG 124

Search completed: April 14, 2006, 17:34:26
 Job time : 7.7465 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: April 14, 2006, 17:04:10 ; Search time 36.1588 Seconds

(without alignments)
1912.171 Million cell updates/sec

Title: US-10-620-246-88

Perfect score: 479
Sequence: 1 MEKSHDPIADIGTQVSDN.....VQDVARTYSQIDGAGVFA 98

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt_05.80.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	479	100.0	98	Q7TVG4_MYCBO	Q7TVG4 mycobacteri
2	479	100.0	99	Q79F93_MYCTU	Q79F93 mycobacteri
3	479	100.0	112	Q8VIR8_MYCTU	Q8VIR8 mycobacteri
4	231	48.2	111	Q7TVS7_MYCBO	Q7TVS7 mycobacteri
5	231	48.2	111	Q79F91_MYCTU	Q79F91 mycobacteri
6	231	48.2	123	Q7D4Y1_MYCTU	Q7D4Y1 mycobacteri
7	102.5	21.4	102	Q73TD7_MYCPA	Q73TD7 mycobacteri
8	102.5	21.4	102	Q7U2D0_MYCBO	Q7U2D0 mycobacteri
9	102.5	21.4	102	Q7DA36_MYCTU	Q7DA36 mycobacteri
10	92.5	19.3	639	Q79F93_MYCTU	Q79F93 mycobacteri
11	92.5	19.3	650	Q8VUW1_MYCTU	Q8VUW1 mycobacteri
12	91	19.0	4470	Q4KX47_LEISM	Q4KX47 leishmania
13	90.5	18.9	87	Q7T212_MYCBO	Q7T212 mycobacteri
14	90	18.8	588	Q6MXX7_MYCTU	Q6MXX7 mycobacteri
15	90	18.8	626	Q7TMM2_MYCBO	Q7TMM2 mycobacteri
16	90	18.8	628	Q8VUJ9_MYCTU	Q8VUJ9 mycobacteri
17	87.5	18.3	606	Q79FV3_MYCTU	Q79FV3 mycobacteri
18	87.5	18.3	608	Q7U125_MYCBO	Q7U125 mycobacteri
19	87.5	18.3	609	Q8VKC5_MYCTU	Q8VKC5 mycobacteri
20	87	18.2	102	PE15_MYCBO	PE15 mycobacteri
21	87	18.2	102	PE15_MYCBO	PE15 mycobacteri
22	86.5	18.1	783	Q7D9C6_MYCTU	Q7D9C6 mycobacteri
23	86.5	18.1	783	Q79FW8_MYCTU	Q79FW8 mycobacteri
24	86.5	18.1	992	Q7TWB8_MYCBO	Q7TWB8 mycobacteri
25	86.5	18.1	1217	Q8V1Y9_MYCTU	Q8V1Y9 mycobacteri
26	86.5	18.1	1460	Q7TMC3_MYCBO	Q7TMC3 mycobacteri
27	86.5	18.1	1489	Q6MMW6_MYCTU	Q6MMW6 mycobacteri
28	86.5	18.1	1715	Q8V1Z0_MYCTU	Q8V1Z0 mycobacteri
29	86.5	18.1	1901	PE54_MYCTU	PE54 mycobacteri
30	86	18.0	233	Q6NAD9_RHOA	Q6NAD9 rhodopsin
31	86	18.0	4454	Q4FYB6_LEIMA	Q4FYB6 leishmania

32	85.5	17.8	543	1	PG44_MYCTU	Q50630 mycobacteri
33	85.5	17.8	546	2	Q7TY98_MYCBO	Q7TY98 mycobacteri
34	85.5	17.8	801	1	PG10_MYCTU	Q53810 mycobacteri
35	85.5	17.8	909	2	Q7UID3_MYCBO	Q7UID3 mycobacteri
36	85	17.7	452	2	Q7U049_MYCBO	Q7U049 mycobacteri
37	85	17.7	487	2	Q79G08_MYCTU	Q79G08 mycobacteri
38	85	17.7	538	2	Q7U2T0_MYCBO	Q7U2T0 mycobacteri
39	85	17.7	561	2	Q8VKR1_MYCTU	Q8VKR1 mycobacteri
40	85	17.7	691	2	Q89Y59_BRADJA	Q89Y59 bradyrhizob
41	83.5	17.4	797	2	Q7UID4_MYCBO	Q7UID4 mycobacteri
42	83	17.3	501	2	Q6V217_SENTR	Q6V217 salmonella
43	83	17.3	501	2	Q6V218_SENTR	Q6V218 salmonella
44	83	17.3	576	2	P71664_MYCTU	P71664 mycobacteri
45	82	17.1	397	2	Q6N223_RHOA	Q6N223 rhodopsin

ALIGNMENTS

RESULT 1
Q7TVG4_MYCBO
ID Q7TVG4_MYCBO PRELIMINARY; PRT; 98 AA.

AC Q7TVG4_1
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE PR FAMILY-LIKE PROTEIN.

GN Name=PE15; OrderedLocNames=MB3902;

OS Mycobacterium bovis.

OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;

OC Corynebacteriaceae; Mycobacteriaceae; Mycobacterium;

OC Mycobacterium tuberculosis complex.

OK NCBI_TaxID=1765;

RN [1]

RC NUCLEOTIDE SEQUENCE.

RX MEDLINE=22709107; PubMed=12788972; DOI=10.1073/pnas.1130426100;

RA Garnier T., Bigliamer K., Camus J.-C., Medina N., Mansoor H.,

RA Pryor M., Duthoy S., Grondin S., Lacroix C., Monsemp C., Simon S.,

RA Harris B., Actin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,

RA Parthill J., Barrell B.G., Cole S.T., Gordon S.V., Hewison R.G.,

RT "The complete genome sequence of Mycobacterium bovis.";

RL Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882(2003).

DR EMBL; BX248347; CAD96088.1; -; Genomic_DNA.

Complete proteome.

SEQUENCE 98 AA; 9799 MW; DAEB998A8BFFB7 CRC64;

Query Match 100.0%; Score 479; DB 2; Length 98;
Best Local Similarity 100.0%; Pred. No. 1.6e-36;
Matches 98; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 MEKSHDPIADIGTQVSDN...HGTAGSTALTSVTGLVPAGAEVSAQAATFTSISIQ 60
Db 1 MEKSHDPIADIGTQVSDN...HGTAGSTALTSVTGLVPAGAEVSAQAATFTSISIQ 60

61 LIASNSAQDQIHRAGAVQDVARTYSQIDGAGVFA 98
61 LIASNSAQDQIHRAGAVQDVARTYSQIDGAGVFA 98

RESULT 2

Q79F93_MYCTU
ID Q79F93_MYCTU PRELIMINARY; PRT; 99 AA.

AC Q79F93_1

DT 05-JUL-2004 (TrEMBLrel. 27, Created)

DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)

DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)

DE PR FAMILY-RELATED PROTEIN.

GN Name=PE15; OrderedLocNames=Rv3872;

OS Mycobacterium tuberculosis.

OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;

OC Corynebacteriaceae; Mycobacteriaceae; Mycobacterium;

OC Mycobacterium tuberculosis complex.

RA Sulston J.E., Taylor K., Whited S., Barrell B.G.;
RT "Disciphering the biology of *Mycobacterium tuberculosis* from the
RT complete genome sequence."
RL Nature 393:537-544 (1998).
DR EMBL; BX042584; CABS5636.1; -; Genomic_DNA.
DR TubercuList; RV3746c; -.
KM Complete proteome.
SQ SEQUENCE 111 AA; 11496 MW; 0F03FBCL12822E996 CRC64;

Query Match	48.2%	Score 231	DB 23	Length 111
Best Local Similarity	52.1%	Pred. No. 1.4e-13		
Matches 49	Conservative 15	Mismatches 30	Indels 0	Gaps 0

QY	Db
1	1
MEKMSHDP	1
PIAADTCTGYSDNALHGVTAGTALTSVGL	1
PAGAD EVSACATATPTEGIC	60
1	1
MOSMFPPADVAIDIGSQVNNNA	1
FOGLAQGAVAWMTSLSTLLPAGABREVSAMVTAFTTAATG	60
61	61
LLASNAASQODLHRAGAEVDDVARTSYQIDDCAA	94
61	61
LIALNQAQSEELRKAGEVFTALIRMTSDADVRA	94

RESULT 6	
Q7D4Y1_MYCTU	
ID	
Q7D4Y1_MYCTU	PRELIMINARY;
PRT;	123 AA

DT 05-JUL-2004 (TRMBLrel. 27, Created)
 DT 05-JUL-2004 (TRMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TRMBLrel. 27, Last annotation update)
 DE PR family protein.
 GN OrderedLocusNames=MT3854;
 OS Mycobacterium tuberculosis.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales
 OC Corynebacterineae; Mycobacteriaceae; Mycobacterium;
 OC Mycobacterium tuberculosis complex.
 NC NCBI_taxid=1773;

RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CDC 1551 / Oehkesh;
RX MEDLINE=22206494; PubMed=12218036;
RX DOI=10.1128/JB.184.19.5479-5490.2002;
RA Pleischmann R.D., Allard D., Bisen J.A.,
RA Peterson J.D., Deboy R.T., Dodson R.J.,
RA Hickey E.K., Kolonay J.F., Nelson W.C.,
RA Salzberg S.L., Delcher A., Uterback T.R.,
RA Gill J., Mikula A., Blehni W., Jacobs W.R. Jr.,
RA Fraser C.M.;
RT "Whole-genome comparison of *Mycobacterium tuberculosis* clinical and laboratory strains.";
RT J. Bacteriol. 184:5479-5490(2002).
DR EMBL: A8000516; AAK48218.1; -, Genomic_DNA.
misc. info

Query Match	48.2%	Score 231	DB 2	length 123
Best Local Similarity	52.1%	Pred. No.	1.5e-13	
Matches 49	Conservative	15	Mismatches	30
			Indels	0
			Gaps	0

Qy	1	MERKSHPIADIGTGVSDNALHGVNAGS ⁺ PALTSVGLVPAGADEVSACATAPTEBGIC	6
	:	: : : : : :	:
	:	: : : : : :	:
	:	: : : : : :	:
	:	: : : : : :	:
Db	13	MQGMSFPDPAVDIGSGCVNNNAFQGLQAGAVAMWSTSLILPAGABRYSAWMTAFITTAANS	72
Qy	61	LIASNSAQOQLRAGEAVIDVARTEYQLIDGAA	94
	:	: : : : : :	:
	:	: : : : : :	:
	:	: : : : : :	:
	:	: : : : : :	:
Db	73	LIALNQAOEELRKRAGEFTTALAANSDADVRAA	106

Q73TD7_MYCPA	RESULT 7
ID Q73TD7_MYCPA	PRELIMINARY; PRT; 102 AA.
AC Q73TD7	
DT 05-JUL-2004	(TRIMBAbrel. 27, Created)
DT 05-JUL-2004	(TRIMBAbrel. 27, Last sequence update)

D7 05-JUN-2004 (TRIMBLrel. 27, Last annotation update)
DE Hypothetical protein.
GN OrderedLocNames=MAP3781;
OS Mycobacterium paratuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium;
OC Mycobacterium avium complex (MAC).
NCBI_TaxID=1770;

NS
RP
RC
RA
RL
DR
EMBL; AS017240; AAS0631.1; -; Genomic DNA.
Interpro: IPR000084; PE region N.

Query Match	21.4%	Score 102.5;	DB 2;	Length 102;
Best Local Similarity	31.5%	Pred. No. 0.089;		
Matches 28;	Conservative 14;	Mismatches 42;	Indels 5;	Gaps 1

QY 14 GTGVSNALHLGVT-----AGSTLTAVTGTVPGVGADEVSAQAATATSTSGI QLLANNSA 68
 Db 9 GLAATSAAYEALFARLAAAHHAAAPRITNVPPRADPVELQTALGRSAQGGHSAVAAGG 66
 QY 69 QDQHLRAGERVQDVATTSQIDDGAGVF 97
 Db 69 VEEELGRAGVGNGGAGSYLAGDTAAATATP 97

RESULT 8
Q7U2D0 MYCBO
ID Q7U2D0 MYCBO PRELIMINARY; PRT; 102 AA

DT 01-OCT-2003 (TReMBLrel. 25, Created)
 DT 01-OCT-2003 (TReMBLrel. 25, Last sequence update)
 DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
 DE PE FAMILY PROTEIN.
 GN Name=PE5; OrderedLocustNames=Mb0293;
 OS Mycobacterium Dovic.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales
 OC Corynebacterineae; Mycobacteriaceae; Mycobacterium;
 OC Mycobacterium tuberculosis complex.
 NCBI_TaxID=1765;
 OX

RP NUCLEOTIDE SEQUENCE.
RC STRAIN=AF2122/97;
MEDLINE=22709107; PubMed=12786972; DOI=10.1073/pnas.1130426100;
RA Garner T., Eglymeler K., Camus J.-C., Medina N., Mansoor H.,
RA Pryor M., Duthoy S., Grondin S., Lacroix C., Monsempc C., Simon S.,
RA Harris B., Atlin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,
RA Parkhill J., Barrrell B.G., Cole S.T., Gordon S.V., Hewinson R.G.;
RA "The complete genome sequence of *Mycobacterium bovis*."
RT Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882 (2003).
DR EMBL: BX24835; CAD93157.1; -; Genomic_DNA.
DR InterPro: IPR000084; Pfam_region_N.

Query Match	21.4%	Score 102.5	DB 2	length 102
Best Local Similarity	29.2%	Pred. No. 0.089		
Matches 26; Conservative	17;	Mismatches	41;	Indels 5; Gaps 1;

QY 1 GGVGSNNAALHGVT-----AGSTATSVGLVGPAGDEVSQAQAATFTSGIQLAASNSA 68
Db 9 GLAAASAAVEALTTLRLAAAHASAPVITAVNPPADPVSQTAPAGFSAGCVENAVVTLAG 68
QY 69 QDQDLHAGAAVODVARTTYSQIDDDGAAGVF 97
Db 69 VTELGRAGVGVESGASVYLAGDAAAATY 97

"Whole-genome comparison of Mycobacterium tuberculosis clinical and laboratory strains.";
RT J. Bacteriol. 184:5479-5490(2002).
DR EMBL/ AEO00516; AAK46124.1; -; Genomic_DNA.
DR TIGR/ MT1853; -;
DR InterPro; IPR002952; Eggshell.
DR InterPro; IPR000884; PE_region_N.
DR InterPro; IPR002173; PfkB.
DR Pfam; PF00934; PE; 1.
DR PRINTS; PR01228; EGGSHLL.
DR Prodom; PD001223; PE_region_N; 1.
DR PROSITE; PS00583; PFKB_KINASES_1; UNKNOWN_1.
SQ SEQUENCE 650 AA; 56541 MW; C5D5216FD7CCD87 CRC64;

Query Match 19.3%; Score 92.5; DB 2; Length 650;
Best Local Similarity 27.1%; Pred. No. 5.4;
Matches 32; Conservative 15; Mismatches 32; Indels 39; Gaps 4;

QY 11 ADIGTQVSDNALHGVTAAGTALTSVTGLVPAAGADEVSAQAATATFTSG-----IQLLASNLSA 58
DB 32 ATIGSAIS-----RANAELVPTLALPAGADVSAALALPATRGQAYQELSAHAVA 84

QY 59 -----IQLLASNLSA-----ODLHRAGEAVQVARTYSO-----IDDGAGV 96
DB 85 FHEQFVQVMSAGAAQVYASAPANSSPIQVGGTALDAINSVQTLTGRLPGANGV 142

RESULT 12
QAKKA7 LEIME
ID QAKKA7_LEIME PRELIMINARY; PRT; 4470 AA.
AC QAKKA7;
DT 13-SEP-2005 (TREMBLrel. 31, Created)
DT 13-SEP-2005 (TREMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TREMBLrel. 31, Last annotation update)
DB Cytoplasmic dynein heavy chain 2.2.
OS Leishmania mexicana.
OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCBI_TaxID=5665;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MYC/BZ/62/M379;
RA Ahlamo C., Forney J.D., Asai D.J., Lebowitz J.H.;
RT "Cytoplasmic dynein 2 isoform is required for flagella assembly in Leishmania mexicana."
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY671898; AAU93603.1; -; Genomic DNA.
SQ SEQUENCE 4470 AA; 500686 MW; 81CF1DAF0226A8D2 CRC64;

Query Match 19.0%; Score 91; DB 2; Length 4470;
Best Local Similarity 30.3%; Pred. No. 58;
Matches 33; Conservative 17; Mismatches 45; Indels 14; Gaps 5;

QY 1 MEKSH-DRPADIGTQVSDNALHGVTAAGTALTSVTGLVPAAGADEVSAQAATATFTSG 58
DB 1584 LEMUGSQNSVTVQVHLKGLFMGIHATFSSDN-TSITHMISADGEVALRRPVATGSD 1642

QY 59 IO--ILASNASADPOLH-----RAGBAVQVARTYS-SCIDGAGV 96
DB 1643 VEMVLLADNASMODTTLHLLTSCVQDPATTKOSVLTYPQIILQVAGCV 1691

RESULT 13
Q7TZ12 MYCBO
ID Q7TZ12_MYCBO PRELIMINARY; PRT; 87 AA.
AC Q7TZ12;
DT 01-OCT-2003 (TREMBLrel. 25, Created)
DT 01-OCT-2003 (TREMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DB PE-PGRS FAMILY PROTEIN.
GN Name=PE_PGRS32a; OrderedlocusNames=Db1832c;
OS Mycobacterium bovis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales; Corynebacteriaceae; Mycobacteriaceae; Mycobacterium;

OC Mycobacterium tuberculosis complex.
OX NCBI_TaxID=1765;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=AF2122/97;
EX MEDLINE=22709107; PubMed=12788972; DOI=10.1073/pnas.1130426100;
RA Garnier T., Biglmeier K., Camus J.-C., Medina N., Mansoor H.,
RA Pryor M., Duthey S., Grondin S., Lacroix C., Monempe C., Simon S.,
RA Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,
RA Parkhill J., Barrrell B.G., Cole S.T., Gordon S.V., Hewison R.G.;
RT "The complete genome sequence of Mycobacterium bovis."
RL Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882(2003).
DR EMBL; BX248340; CAD94535.1; -; Genomic_DNA.
DR InterPro; IPR000084; PE_region_N.
DR Pfam; PF00934; PE; 1.
DR Prodom; PD001223; PE_region_N; 1.
KM Complete proteome.
SQ SEQUENCE 87 AA; 8824 MW; 3EAP1686EAC68C3 CRC64;

Query Match 18.9%; Score 90.5; DB 2; Length 87;
Best Local Similarity 39.2%; Pred. No. 0.96;
Matches 29; Conservative 10; Mismatches 22; Indels 13; Gaps 3;

QY 11 ADIGTQVSDNALHGVTAAGTALTSVTGLVPAAGADEVSAQAATATFTSG--IQLLASNLSA 68
DB 21 ATIGSAIS-----RANAELVPTLALPAGADVSAALALPATRGQAYQELSAHAVA 73

QY 69 QD-----OLHRAGEA 78
DB 74 FHEQFVQVMSAGAA 87

RESULT 14
Q6MXK7 MYCTU
ID Q6MXK7_MYCTU PRELIMINARY; PRT; 588 AA.
AC Q6MXK7;
DT 05-JUL-2004 (TREMBLrel. 27, Created)
DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
DB PE-PGRS FAMILY PROTEIN.
GN Name=PE_PGRS51; OrderedlocusNames=Rv3367;
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales; Corynebacteriaceae; Mycobacteriaceae; Mycobacterium;
OC Mycobacterium tuberculosis complex.
OX NCBI_TaxID=1773;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=H37Rv;
RX MEDLINE=98295987; PubMed=9634230; DOI=10.1038/31159;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Church C.M.,
RA Harris B., Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III,
RA Tekale F., Badcock K., Basham D., Brown D., Chillingworth T.,
RA Connor R., Davies R.M., Devlin K., Fellwell T., Gentles S., Hamlin N.,
RA Holroyd S., Hornsby T., Jagels K., Krogh A., McLean J., Moule S.,
RA Murphy L.D., Oliver S., Osborne J., Quail M.A., Rajandream M.A.,
RA Rogers J., Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sultson J.E., Taylor K., Whitehead S., Barrrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence."
RL Nature 393:537-544(1998).
DR EMBL; BX842582; CA855590.1; -; Genomic_DNA.
DR Tuberculist; RV3367; -;
DR InterPro; IPR002952; Eggshell.
DR InterPro; IPR000084; PE_region_N.
DR InterPro; IPR002173; PfkB.
DR Pfam; PF00934; PE; 1.
DR PRINTS; PR01228; EGGSHLL.
DR PROSITE; PS00583; PFKB_KINASES_1; UNKNOWN_1.
KM Complete proteome.
SQ SEQUENCE 588 AA; 49708 MW; 067B84097F81DAF1 CRC64;

Query Match 18.8%; Score 90; DB 2; Length 588;

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 14, 2006, 17:32:59 ; Search time 9.0618 Seconds

(without alignments)
894.107 Million cell updates/sec

Title: US-10-620-246-88

Perfect score: 479

Sequence: 1 MEMSHDPIADIGTGVSDN.....VQDVARTYSQIDGAGVFA 98

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: Issued_Patents_AA.*

2: /cgn2_6/ptodata/1/1aa/5.COMB.pep.*

3: /cgn2_6/ptodata/1/1aa/6.COMB.pep.*

4: /cgn2_6/ptodata/1/1aa/H.COMB.pep.*

5: /cgn2_6/ptodata/1/1aa/PCBUS.COMB.pep.*

6: /cgn2_6/ptodata/1/1aa/BACKFILE1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	479	100.0	98	US-09-050-739-88	Sequence 88, Appl
2	74.5	15.6	99	US-09-073-009-32	Sequence 32, Appl
3	74.5	15.6	99	US-09-073-009-32	Sequence 32, Appl
4	74.5	15.6	99	US-09-073-010-32	Sequence 32, Appl
5	74.5	15.6	99	US-09-073-010-32	Sequence 32, Appl
6	74.5	15.6	212	US-09-125-619-32	Sequence 32, Appl
7	74.5	15.6	212	US-10-222-566-32	Sequence 32, Appl
8	74.5	15.6	212	US-10-143-024A-32	Sequence 32, Appl
9	74.5	15.6	212	US-10-222-162-32	Sequence 32, Appl
10	74.5	15.6	213	US-09-125-619-39	Sequence 39, Appl
11	74.5	15.6	213	US-10-222-566-39	Sequence 39, Appl
12	74.5	15.6	213	US-10-143-024A-39	Sequence 39, Appl
13	74.5	15.6	213	US-10-222-162-39	Sequence 39, Appl
14	74.5	15.6	213	US-10-222-162-39	Sequence 39, Appl
15	74.5	15.6	505	US-09-603-208A-292	Sequence 292, App
16	74.5	15.6	505	US-09-605-703B-2330	Sequence 2330, App
17	72	15.0	966	US-08-571-758-2	Sequence 2, Appl
18	72	15.0	966	US-08-909-984A-2	Sequence 2, Appl
19	71.5	14.9	620	US-08-909-983-2	Sequence 2, Appl
20	71.5	14.9	713	US-09-902-540-12224	Sequence 12224, A
21	71.5	14.9	723	US-09-540-236-3595	Sequence 3595, Ap
22	71.5	14.9	1001	US-08-548-159-5	Sequence 5, Appl
23	71.5	14.8	438	US-08-884-569A-2	Sequence 2, Appl
24	70	14.6	717	US-09-252-991A-20127	Sequence 20127, A
25	69.5	14.5	892	US-08-435-925C-2	Sequence 2, Appl
26	69.5	14.5	892	US-09-252-991A-32442	Sequence 32442, A
27	69.5	14.5	892	US-09-336-447A-5	Sequence 5, Appl

28	69	14.4	210	2	US-09-710-279-2864	Sequence 2864, Ap
29	69	14.4	210	2	US-09-710-279-3320	Sequence 3320, Ap
30	69	14.4	259	2	US-09-252-991A-30663	Sequence 30663, A
31	69	14.4	288	2	US-09-134-001C-4969	Sequence 4969, Ap
32	69	14.4	1938	1	US-09-949-016-6609	Sequence 6609, Ap
33	69	14.4	2035	1	US-08-046-585-5	Sequence 5, Appl
34	69	14.4	2035	1	US-08-393-703-5	Sequence 5, Appl
35	69	14.4	2035	4	PCT-US93-11721-5	Sequence 5, Appl
36	68.5	14.3	393	2	US-09-252-991A-28810	Sequence 28810, A
37	68	14.2	138	2	US-09-252-991A-17846	Sequence 17846, A
38	68	14.2	141	2	US-09-252-991A-26729	Sequence 26729, A
39	68	14.2	287	2	US-09-107-532A-4761	Sequence 4761, Ap
40	68	14.2	370	2	US-09-902-540-13987	Sequence 13987, A
41	68	14.2	290	2	US-09-252-991A-32469	Sequence 32469, A
42	68	14.2	2680	2	US-09-489-039A-7973	Sequence 7973, Ap
43	67.5	14.1	741	2	US-09-902-540-15369	Sequence 15369, A
44	67	14.0	180	6	5273901-7	Patent No. 5273901
45	67	14.0	180	6	5482709-6	Patent No. 5482709

ALIGNMENTS

```

RESULT 1
US-09-050-739-88
; Sequence 88, Application US/09050739
; Patent No. 6641814
; GENERAL INFORMATION:
; APPLICANT: ANDERSEN, Peter
; APPLICANT: NIELSEN, Rikke
; APPLICANT: OETTINGER, Thomas
; APPLICANT: RASMUSSEN, Peter Birk
; APPLICANT: ROSENKRANDS, Ida
; APPLICANT: WELDLING, Karin
; APPLICANT: FLORIO, Walter
; TITLE OR INVENTION: NUCLEIC ACIDS FRAGMENTS AND POLYPEPTIDE FRAGMENTS
; FILE REFERENCE: 670001-2002.1
; CURRENT APPLICATION NUMBER: US/09/050,739
; CURRENT FILING DATE: 1998-03-30
; EARLIER APPLICATION NUMBER: 0376/97
; EARLIER FILING DATE: 1997-04-02
; EARLIER APPLICATION NUMBER: 1277/97
; EARLIER FILING DATE: 1997-11-10
; EARLIER APPLICATION NUMBER: 60/044,624
; EARLIER FILING DATE: 1997-04-18
; EARLIER APPLICATION NUMBER: 60/070,488
; EARLIER FILING DATE: 1997-01-05
; NUMBER OF SEQ ID NOS: 173
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 88
; LENGTH: 98
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-09-050-739-88

Query Match      100.0%; Score 479; DB 2; Length 98;
Best Local Similarity 100.0%; Pred. No. 5.6e-50;
Matches 98; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MEMSHDPIADIGTGVSDNALHGVTRAGTATSVTGLVPAAGADEVSAQAATAFTSGIQ 60
      |||
Db      1 MEMSHDPIADIGTGVSDNALHGVTRAGTATSVTGLVPAAGADEVSAQAATAFTSGIQ 60

QY      61 ILASNSAQOQLRAGEAVQDVARTYSQIDGAGVFA 98
      |||
Db      61 ILASNSAQOQLRAGEAVQDVARTYSQIDGAGVFA 98

RESULT 2
US-09-073-009-32
; Sequence 32, Application US/09073009
; Patent No. 655653

```



```

; GENERAL INFORMATION:
; APPLICANT: Alderson, Mark
; APPLICANT: Dillon, David C.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Campos-Neto, Antonio
; TITLE OF INVENTION: COMPOUNDS AND DIAGNOSIS OF
; NUMBER OF SEQUENCES: 144
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY
; STREET: 6300 Columbia Center, 701 Fifth Ave.
; CITY: Seattle
; STATE: Washington
; COUNTRY: US
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/073,009
; FILING DATE: 05-MAY-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.441C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-622-4900
; TELEFAX: 206-682-6031
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 99 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: Mycobacterium tuberculosis
;
US-09-073-009-32
;
Query Match 15.6%; Score 74.5; DB 2; Length 99;
Best Local Similarity 39.7%; Pred. No. 0.21;
Matches 23; Conservative 8; Mismatches 18; Indels 9; Gaps 3;
;
Cy 13 IGTQSDNALHGVTAGTALTSTGLVPAGADEVSAQAATFTSEG--IQLLASNSA 68
Db 21 IGTMM--NAQNMAAAAPT-----TGVVPAADAEVSALTAQFAAHQMGTVSQAQAA 71
;
; RESULT 3
; US-09-073-009-33
; Sequence 33, Application US/09073009
; Patent No. 6555653
; GENERAL INFORMATION:
; APPLICANT: Alderson, Mark
; APPLICANT: Dillon, David C.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Campos-Neto, Antonio
; TITLE OF INVENTION: COMPOUNDS AND DIAGNOSIS OF
; NUMBER OF SEQUENCES: 144
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY
; STREET: 6300 Columbia Center, 701 Fifth Ave.
; CITY: Seattle
; STATE: Washington
; COUNTRY: US
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
;
TUBERCULOSIS AND ME
```

```

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/073,009
; FILING DATE: 05-MAY-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.441C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-622-4900
; TELEFAX: 206-682-6031
; INFORMATION FOR SEQ ID NO: 33:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 99 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: Mycobacterium tuberculosis
;
US-09-073-009-33
;
Query Match 15.6%; Score 74.5; DB 2; Length 99;
Best Local Similarity 39.7%; Pred. No. 0.21;
Matches 23; Conservative 8; Mismatches 18; Indels 9; Gaps 3;
;
Cy 13 IGTQSDNALHGVTAGTALTSTGLVPAGADEVSAQAATFTSEG--IQLLASNSA 68
Db 21 IGTMM--NAQNMAAAAPT-----TGVVPAADAEVSALTAQFAAHQMGTVSQAQAA 71
;
; RESULT 4
; US-09-073-010-32
; Sequence 32, Application US/09073010
; Patent No. 6613881
; GENERAL INFORMATION:
; APPLICANT: Alderson, Mark
; APPLICANT: Dillon, David C.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Campos-Neto, Antonio
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS OF TUBERCULOSIS AND
; NUMBER OF SEQUENCES: 144
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY
; STREET: 6300 Columbia Center, 701 Fifth Ave.
; CITY: Seattle
; STATE: Washington
; COUNTRY: US
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/073,010
; FILING DATE: 05-MAY-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.440C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-622-4900
; TELEFAX: 206-682-6031
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 99 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
;
TUBERCULOSIS AND ME
```



```

; APPLICANT: JING-REN, ZHANG
; APPLICANT: HARDHAM, JOHN M.
; APPLICANT: HOWELL, JERRILYN K.
; APPLICANT: BARBOUR, ALAN G.
; APPLICANT: WEINSTOCK, GEORGE M.
; TITLE OF INVENTION: VMP-LIKE SEQUENCES OF PATHOGENIC BORRELLIA
; FILE REFERENCE: UTSH:234USD1
; CURRENT APPLICATION NUMBER: US/10/143,024A
; PRIOR FILING DATE: 2002-08-23
; PRIOR APPLICATION NUMBER: 09/125,619
; PRIOR FILING DATE: 1999-01-27
; PRIOR APPLICATION NUMBER: PCT/US97/02952
; PRIOR FILING DATE: 1997-02-20
; PRIOR APPLICATION NUMBER: 60/012,028
; PRIOR FILING DATE: 1996-02-21
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 32
; LENGTH: 212
; TYPE: PRT
; ORGANISM: Borrelia burgdorferi
US-10-143-024A-32
```

```
Query Match      15.6%; Score 74.5; DB 2; Length 212;
Best Local Similarity 26.7%; Pred. No. 0.62;
Matches 20; Conservative 21; Mismatches 29; Indels 5; Gaps 2;
```

```
QY 24 GTTAGSTALTSVGLVPAGADSVSAQAATAFTSEGIQLASNSAODQLHRA---GEAVQ 80
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 39 GASSGTDALGEV--VANAGAAKAVADKASVTGIAKGIKEIVEAAGSGSEKLVAAATGESNK 96
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 81 DVARTYSQIDDDGAG 95
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 97 GAGKLFKGVDDAHAG 111
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
```

```
RESULT 9
US-10-222-162-32
; Sequence 32, Application US/10222162
; Patent No. 6878816
; GENERAL INFORMATION:
; APPLICANT: NORRIS, STEVEN J.
; APPLICANT: JING-REN, ZHANG
; APPLICANT: HARDHAM, JOHN M.
; APPLICANT: HOWELL, JERRILYN K.
; APPLICANT: BARBOUR, ALAN G.
; APPLICANT: WEINSTOCK, GEORGE M.
; TITLE OF INVENTION: VMP-LIKE SEQUENCES OF PATHOGENIC BORRELLIA
; FILE REFERENCE: UTSH:234USD4
; CURRENT APPLICATION NUMBER: US/10/222,162
; CURRENT FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: 09/125,619
; PRIOR FILING DATE: 1999-01-27
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 32
; LENGTH: 212
; TYPE: PRT
; ORGANISM: Borrelia burgdorferi
US-10-222-162-32
```

```
Query Match      15.6%; Score 74.5; DB 2; Length 212;
Best Local Similarity 26.7%; Pred. No. 0.62;
Matches 20; Conservative 21; Mismatches 29; Indels 5; Gaps 2;
```

```
QY 24 GTTAGSTALTSVGLVPAGADSVSAQAATAFTSEGIQLASNSAODQLHRA---GEAVQ 80
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 39 GASSGTDALGEV--VANAGAAKAVADKASVTGIAKGIKEIVEAAGSGSEKLVAAATGESNK 96
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 81 DVARTYSQIDDDGAG 95
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 97 GAGKLFKGVDDAHAG 111
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
```

```

; APPLICANT: JING-REN, ZHANG
; APPLICANT: HARDHAM, JOHN M.
; APPLICANT: HOWELL, JERRILYN K.
; APPLICANT: BARBOUR, ALAN G.
; APPLICANT: WEINSTOCK, GEORGE M.
; TITLE OF INVENTION: VMP-LIKE SEQUENCES OF PATHOGENIC BORRELLIA
; FILE REFERENCE: UTSH:234
; CURRENT APPLICATION NUMBER: US/09/125,619
; CURRENT FILING DATE: 1999-01-27
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 39
; LENGTH: 213
; TYPE: PRT
; ORGANISM: Borrelia burgdorferi
US-09-125-619-39
```

```
Query Match      15.6%; Score 74.5; DB 2; Length 213;
Best Local Similarity 26.7%; Pred. No. 0.63;
Matches 20; Conservative 21; Mismatches 29; Indels 5; Gaps 2;
```

```
QY 24 GTTAGSTALTSVGLVPAGADSVSAQAATAFTSEGIQLASNSAODQLHRA---GEAVQ 80
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 39 GASSGTDALGEV--VANAGAAKAVADKASVTGIAKGIKEIVEAAGSGSEKLVAAATGESNK 96
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 81 DVARTYSQIDDDGAG 95
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 97 GAGKLFKGVDDAHAG 111
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
```

```
RESULT 11
US-10-222-566-39
; Sequence 39, Application US/10222566
; Patent No. 6719983
; GENERAL INFORMATION:
; APPLICANT: NORRIS, STEVEN J.
; APPLICANT: JING-REN, ZHANG
; APPLICANT: HARDHAM, JOHN M.
; APPLICANT: HOWELL, JERRILYN K.
; APPLICANT: BARBOUR, ALAN G.
; APPLICANT: WEINSTOCK, GEORGE M.
; TITLE OF INVENTION: VMP-LIKE SEQUENCES OF PATHOGENIC BORRELLIA
; FILE REFERENCE: UTSH:234USD3
; CURRENT APPLICATION NUMBER: US/10/222,566
; CURRENT FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: 09/125,619
; PRIOR FILING DATE: 1999-01-27
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 39
; LENGTH: 213
; TYPE: PRT
; ORGANISM: Borrelia burgdorferi
US-10-222-566-39
```

```
Query Match      15.6%; Score 74.5; DB 2; Length 213;
Best Local Similarity 26.7%; Pred. No. 0.63;
Matches 20; Conservative 21; Mismatches 29; Indels 5; Gaps 2;
```

```
QY 24 GTTAGSTALTSVGLVPAGADSVSAQAATAFTSEGIQLASNSAODQLHRA---GEAVQ 80
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 39 GASSGTDALGEV--VANAGAAKAVADKASVTGIAKGIKEIVEAAGSGSEKLVAAATGESNK 96
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 81 DVARTYSQIDDDGAG 95
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 97 GAGKLFKGVDDAHAG 111
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
```

RESULT 15
US-09-605-703B-2330
; Sequence 2330, Application US/09605703B


```
; Sequence 88, Application US/09804980
; Publication No. US20030147897A1
; GENERAL INFORMATION:
; APPLICANT: Statens Serum Institut
; APPLICANT: Anderson, Peter
; TITLE OF INVENTION: M. Tuberculosis Antigens
; FILE REFERENCE: 670001-2002.4
; CURRENT APPLICATION NUMBER: US/09/804,980
; CURRENT FILING DATE: 2001-03-12
; NUMBER OF SEQ ID NOS: 257
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 88
; LENGTH: 98
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-09-804-980-88

Query Match      100.0%; Score 479; DB 3; Length 98;
Best Local Similarity 100.0%; Pred. No. 1.4e-42;
Matches 98; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MERSHDPPIADIGTGVSDNALHGVTAGSTALTSVTGLVPAGADEVSQAQATAFTSGIQ 60
DB      1 MERSHDPPIADIGTGVSDNALHGVTAGSTALTSVTGLVPAGADEVSQAQATAFTSGIQ 60

QY      61 LLASNSAODQLHRAGEAVODVARTYSQIDDDGAAGVFA 98
DB      61 LLASNSAODQLHRAGEAVODVARTYSQIDDDGAAGVFA 98

RESULT 3
US-10-138-473-88
; Sequence 88, Application US/10138473
; Publication No. US20030165525A1
; GENERAL INFORMATION:
; APPLICANT: ANDERSEN, Peter
; APPLICANT: WEIDINGH, Karin
; APPLICANT: HANSEN, Christina Veggerby
; APPLICANT: FLORIO, Walter
; APPLICANT: OKKELS, Li Mei Meng
; APPLICANT: SKJORT, Rikke Louise Vinther
; APPLICANT: ROSENKRANDS, Ida
; APPLICANT: RASMUSSEN, Peter Blirk
; TITLE OF INVENTION: TB Diagnostic Based On Antigens From M. tuberculosis
; FILE REFERENCE: 0459-0710P
; CURRENT APPLICATION NUMBER: US/10/138,473
; CURRENT FILING DATE: 2002-05-02
; PRIOR APPLICATION NUMBER: US 10/060,428
; PRIOR FILING DATE: 2002-01-29
; PRIOR APPLICATION NUMBER: US 09/415,884
; PRIOR FILING DATE: 1999-10-08
; PRIOR APPLICATION NUMBER: US 60/116,673
; PRIOR FILING DATE: 1999-01-21
; PRIOR APPLICATION NUMBER: DK 1998 01281
; PRIOR FILING DATE: 1998-10-18
; PRIOR APPLICATION NUMBER: US 60/070,488
; PRIOR FILING DATE: 1998-01-05
; PRIOR APPLICATION NUMBER: DK 1997 01277
; PRIOR FILING DATE: 1997-11-10
; PRIOR APPLICATION NUMBER: US 60/044,624
; PRIOR FILING DATE: 1997-04-18
; PRIOR APPLICATION NUMBER: DK 1997 00376
; PRIOR FILING DATE: 1997-04-02
; NUMBER OF SEQ ID NOS: 174
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 88
; LENGTH: 98
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-10-138-473-88

Query Match      100.0%; Score 479; DB 4; Length 98;
Best Local Similarity 100.0%; Pred. No. 1.4e-42;
```

```
Matches 98; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MERSHDPPIADIGTGVSDNALHGVTAGSTALTSVTGLVPAGADEVSQAQATAFTSGIQ 60
DB      1 MERSHDPPIADIGTGVSDNALHGVTAGSTALTSVTGLVPAGADEVSQAQATAFTSGIQ 60

QY      61 LLASNSAODQLHRAGEAVODVARTYSQIDDDGAAGVFA 98
DB      61 LLASNSAODQLHRAGEAVODVARTYSQIDDDGAAGVFA 98

RESULT 4
US-10-620-246-88
; Sequence 88, Application US/10620246
; Publication No. US20040115211A1
; GENERAL INFORMATION:
; APPLICANT: ANDERSEN, Peter
; APPLICANT: NIELSEN, Rikke
; APPLICANT: OSTINGER, Thomas
; APPLICANT: RASMUSSEN, Peter Blirk
; APPLICANT: ROSENKRANDS, Ida
; APPLICANT: WEIDINGH, Karin
; APPLICANT: FLORIO, Walter
; TITLE OF INVENTION: NUCLEIC ACIDS FRAGMENTS AND POLYPEPTIDE FRAGMENTS
; FILE REFERENCE: 670001-2002.1A
; CURRENT APPLICATION NUMBER: US/10/620,246
; CURRENT FILING DATE: 2003-07-15
; PRIOR APPLICATION NUMBER: 09/050,739
; PRIOR FILING DATE: 1998-03-30
; PRIOR APPLICATION NUMBER: 0376/97
; PRIOR FILING DATE: 1997-04-02
; PRIOR APPLICATION NUMBER: 1277/97
; PRIOR FILING DATE: 1997-11-10
; PRIOR APPLICATION NUMBER: 60/044,624
; PRIOR FILING DATE: 1997-04-18
; PRIOR APPLICATION NUMBER: 60/070,488
; PRIOR FILING DATE: 1998-01-05
; PRIOR APPLICATION NUMBER: 10/138,473
; PRIOR FILING DATE: 2002-05-02
; PRIOR APPLICATION NUMBER: 09/791,171
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 09/415,884
; PRIOR FILING DATE: 1999-10-08
; PRIOR APPLICATION NUMBER: 60/116,673
; PRIOR FILING DATE: 1999-01-21
; PRIOR APPLICATION NUMBER: 1281/98
; PRIOR FILING DATE: 1998-10-08
; NUMBER OF SEQ ID NOS: 173
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 88
; LENGTH: 98
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-10-620-246-88

Query Match      100.0%; Score 479; DB 4; Length 98;
Best Local Similarity 100.0%; Pred. No. 1.4e-42;
Matches 98; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MERSHDPPIADIGTGVSDNALHGVTAGSTALTSVTGLVPAGADEVSQAQATAFTSGIQ 60
DB      1 MERSHDPPIADIGTGVSDNALHGVTAGSTALTSVTGLVPAGADEVSQAQATAFTSGIQ 60

QY      61 LLASNSAODQLHRAGEAVODVARTYSQIDDDGAAGVFA 98
DB      61 LLASNSAODQLHRAGEAVODVARTYSQIDDDGAAGVFA 98

RESULT 5
US-10-510-021-62
; Sequence 62, Application US/10510021
; Publication No. US20050220811A1
```

```

; GENERAL INFORMATION:
; APPLICANT: Cole, Stewart
; APPLICANT: Pym, Alexander S
; APPLICANT: Brosch, Roland
; APPLICANT: Brodin, Priscille
; APPLICANT: Majlesel, Laleh
; APPLICANT: Demangel, Caroline
; APPLICANT: Leclerc, Claude
; TITLE OF INVENTION: Identification of virulence associated regions RD1 and
; TITLE OF INVENTION: RDS leading to improve vaccine of M. bovis BCG and M.
; TITLE OF INVENTION: microcl
; FILE REFERENCE: D20217
; CURRENT APPLICATION NUMBER: US/10/510,021
; CURRENT FILING DATE: 2004-10-01
; PRIOR APPLICATION NUMBER: PCT/IB03/01789
; PRIOR FILING DATE: 2003-04-01
; PRIOR APPLICATION NUMBER: EP 02/290864
; PRIOR FILING DATE: 2002-04-05
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 62
; LENGTH: 99
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: RV3872-PE35 - PG family-related protein
; US-10-510-021-62

Query Match          100.0%; Score 479; DB 5; Length 99;
Best Local Similarity 100.0%; Pred. No. 1.4e-42;
Matches 98; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEMSHDPIADIGTQVSDNALHGVTAAGTALTSTVGLVPAGADEVSAQAATAFTSGIQ 60
DB 1 MEMSHDPIADIGTQVSDNALHGVTAAGTALTSTVGLVPAGADEVSAQAATAFTSGIQ 60
QY 61 LLASNASADQQLHRAGAVQDVARTYSQIDDGAAQVFA 98
DB 61 LLASNASADQQLHRAGAVQDVARTYSQIDDGAAQVFA 98

RESULT 6
; US-10-282-122A-64609
; Sequence 64609, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl
; APPLICANT: Zykkind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
```

```

; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 7614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 64609
; LENGTH: 639
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
; US-10-282-122A-64609

Query Match          19.3%; Score 92.5; DB 4; Length 639;
Best Local Similarity 27.1%; Pred. No. 0.63;
Matches 32; Conservative 15; Mismatches 32; Indels 39; Gaps 4;

QY 11 ADIGTQVSDNALHGVTAAGTALTSTVGLVPAGADEVSAQAATAFTSGI----- 58
DB 21 ATIGSAIS-----RMAERLVPITLALPAGADVDVSAALPAIPATIGQAVQSLSAHAVA 73
QY 59 -----IQLLASNSA-----QDQLHRAGAVQDVARTYSQ-----IDDGAAQV 96
DB 74 FHRRFQVLMAGAAQVYASAAANSSPLQIVGQRTALDINSFVQTLRPLIGNANQV 131

RESULT 7
; US-10-282-122A-64609
; Sequence 64609, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl
; APPLICANT: Zykkind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
```


US-10-282-1122A-64547
 Sequence 64547, Application US/10282122A
 Publication No. US20040029129A1
 GENERAL INFORMATION:
 APPLICANT: Wang, Liangsu
 APPLICANT: Zamudio, Carlos
 APPLICANT: Malone, Cheryl
 APPLICANT: Haselbeck, Robert
 APPLICANT: Ohlsen, Karl
 APPLICANT: Zyskind, Judith
 APPLICANT: Wall, Daniel
 APPLICANT: Trawick, John
 APPLICANT: Carr, Grant
 APPLICANT: Yamamoto, Robert
 APPLICANT: Forsyth, R.
 APPLICANT: Xu, H.
 TITLE OR INVENTION: Identification of Essential Genes in Microorganisms
 FILE REFERENCE: ERTA 03A
 CURRENT APPLICATION NUMBER: US/10/0282,122A
 CURRENT FILING DATE: 2003-02-20
 PRIOR APPLICATION NUMBER: 60/131,078
 PRIOR FILING DATE: 2000-03-21
 PRIOR APPLICATION NUMBER: 60/206,848
 PRIOR FILING DATE: 2000-05-23
 PRIOR APPLICATION NUMBER: 60/207,727
 PRIOR FILING DATE: 2000-05-26
 PRIOR APPLICATION NUMBER: 60/230,335
 PRIOR FILING DATE: 2000-09-06
 PRIOR APPLICATION NUMBER: 60/230,347
 PRIOR FILING DATE: 2000-09-09
 PRIOR APPLICATION NUMBER: 60/242,578
 PRIOR FILING DATE: 2000-10-23
 PRIOR APPLICATION NUMBER: 60/253,625
 PRIOR FILING DATE: 2000-11-27
 PRIOR APPLICATION NUMBER: 60/257,931
 PRIOR FILING DATE: 2000-12-22
 PRIOR APPLICATION NUMBER: 60/267,636
 PRIOR FILING DATE: 2001-02-09

```

RESULT 12
US-10-282-122A-64514
Sequence 64514, Application US/10282122A
Publication No. US20040029129A1
GENERAL INFORMATION:
APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Onlsen, Karl
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITPA.034A
CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: PatentIn version 3.1
SEQ ID NO 64514
LENGTH: 562
TYPE: PRF
ORGANISM: Mycobacterium tuberculosis
US-10-282-122A-64514

```



```

1  APPLICANT: Zamudio, Carlos
2  APPLICANT: Malone, Cheryl
3  APPLICANT: Haselbeck, Robert
4  APPLICANT: Ohlsen, Kari
5  APPLICANT: Zykend, Judith
6  APPLICANT: Wall, Daniel
7  APPLICANT: Trawick, John
8  APPLICANT: Carr, Grant
9  APPLICANT: Yamamoto, Robert
10 APPLICANT: Forsyth, R.
11 APPLICANT: Xu, H.
12 TITLE OR INVENTION: Identification of Essential Genes in Microorganisms
13 FILE REFERENCE: ELITRA.034A
14 CURRENT APPLICATION NUMBER: US/10/282,122A
15 CURRENT FILING DATE: 2003-02-20
16 PRIOR APPLICATION NUMBER: 60/191,078
17 PRIOR FILING DATE: 2000-03-21
18 PRIOR APPLICATION NUMBER: 60/206,848
19 PRIOR FILING DATE: 2000-05-23
20 PRIOR APPLICATION NUMBER: 60/207,727
21 PRIOR FILING DATE: 2000-05-26
22 PRIOR APPLICATION NUMBER: 60/230,335
23 PRIOR FILING DATE: 2000-09-06
24 PRIOR APPLICATION NUMBER: 60/230,347
25 PRIOR FILING DATE: 2000-09-09
26 PRIOR APPLICATION NUMBER: 60/242,578
27 PRIOR FILING DATE: 2000-10-23
28 PRIOR APPLICATION NUMBER: 60/253,625
29 PRIOR FILING DATE: 2000-11-27
30 PRIOR APPLICATION NUMBER: 60/257,931
31 PRIOR FILING DATE: 2000-12-22
32 PRIOR APPLICATION NUMBER: 60/267,636
33 PRIOR FILING DATE: 2001-02-09
34 PRIOR APPLICATION NUMBER: 60/269,308
35 PRIOR FILING DATE: 2001-02-16
36 Remaining Prior Application data removed - See File Wrapper or PALM.
37 NUMBER OF SEQ ID NOS: 78614
38 SOFTWARE: PatentIn version 3.1
39 SEQ ID NO 64606
40 LENGTH: 914
41 TYPE: PRF
42 ORGANISM: Mycobacterium tuberculosis
43 US-10-282-122A-64606

```

Query Match	16.8%	Score 80.5;	DB 4;	Length 914;
Best Local Similarity	34.5%	Pred. No. 18;		
Matches 20;	Conservative 11;	Mismatches 20;	Indels 7;	Gaps 1;

```

Oy      11 ADIGIQVSDNALHGVTAGSTALTTSVTGIVPAGADEVSAQAAATPTSSEGIQLLSNNSA 68
      ||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      19 ADIGSTTAG-----ANAAAAANTTSILAAAGDEISAIATATLFGAHGRAYQAASAA 69

```

Search completed: April 14, 2006, 18:40:31
Job time : 35.4906 secs

This Page Blank (uspto)


```
; PRIOR FILING DATE: 2004-07-20
; PRIOR APPLICATION NUMBER: IN 173/DEL/2004
; PRIOR FILING DATE: 2004-02-06
; NUMBER OF SEQ ID NOS: 763
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 163
; LENGTH: 606
; TYPE: PRF
; ORGANISM: Mycobacterium tuberculosis H37Rv
US-11-052-554A-163

Query Match      18.3%; Score 87.5; DB 7; Length 606;
Best Local Similarity 32.9%; Pred. No. 0.27;
Matches 27; Conservative 15; Mismatches 29; Indels 11; Gaps 3;

QY 11 ADIGTQVSDNALHGVTAAGTATGTVGTPAGADEVSAQAATAFTSEGIQLLASNSAQ 68
DB 19 AAGTSTLS-----AANAAALAAFTTGTVAAGADEVSAAVASLFSGHAQAVOTLGTQAAA 71

QY 69 QDQHRAGEAVQDVARTYSQID 90
DB 72 FHE--RTQALSTPAGAGVGSAB 91

RESULT 3
US-11-052-554A-157
; Sequence 157, Application US/11052554A
; Publication No. US2005028866A1
; GENERAL INFORMATION:
; APPLICANT: Sachdeva, et al.
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
; FILE REFERENCE: 30853/40359A
; CURRENT APPLICATION NUMBER: US/11/052,554A
; CURRENT FILING DATE: 2005-02-07
; PRIOR APPLICATION NUMBER: US 60/589,227
; PRIOR FILING DATE: 2004-07-20
; PRIOR APPLICATION NUMBER: IN 173/DEL/2004
; PRIOR FILING DATE: 2004-02-06
; NUMBER OF SEQ ID NOS: 763
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 157
; LENGTH: 783
; TYPE: PRF
; ORGANISM: Mycobacterium tuberculosis H37Rv
US-11-052-554A-157

Query Match      18.1%; Score 86.5; DB 7; Length 783;
Best Local Similarity 30.3%; Pred. No. 0.48;
Matches 27; Conservative 14; Mismatches 45; Indels 3; Gaps 1;

QY 13 IGTVQVSDNALHGVTAAGTATGTVGTPAGADEVSAQAATAFTSEGIQLLASNSAQ 69
DB 11 LGSAAIDLAALAGSYLGAADAAATTTGTVAAGADEVSAAIALFSAHGRAYQVASHQAA 70

QY 70 DQHRAGEAVQDVARTYSQIDGAGVFA 98
DB 71 AVHAQFVEALSAAGAGVASABAGAVLA 99

RESULT 4
US-11-052-554A-135
; Sequence 135, Application US/11052554A
; Publication No. US2005028866A1
; GENERAL INFORMATION:
; APPLICANT: Sachdeva, et al.
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
; FILE REFERENCE: 30853/40359A
; CURRENT APPLICATION NUMBER: US/11/052,554A
; CURRENT FILING DATE: 2005-02-07
; PRIOR APPLICATION NUMBER: US 60/589,227
; PRIOR FILING DATE: 2004-07-20
```

```
; PRIOR APPLICATION NUMBER: IN 173/DEL/2004
; PRIOR FILING DATE: 2004-02-06
; NUMBER OF SEQ ID NOS: 763
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 135
; LENGTH: 1901
; TYPE: PRF
; ORGANISM: Mycobacterium tuberculosis H37Rv
US-11-052-554A-135

Query Match      18.1%; Score 86.5; DB 7; Length 1901;
Best Local Similarity 28.9%; Pred. No. 1.6;
Matches 24; Conservative 18; Mismatches 34; Indels 7; Gaps 1;

QY 12 DIGTQVSDNALHGVTAAGTATGTVGTPAGADEVSAQAATAFTSEGIQLLASNSAQDQ 71
DB 20 NAGSSTLS-----AANASASANTTGTVAAGADEVSARIALFGFGLERQALSAQVAAV 72

QY 72 LHRAGEAVQDVARTYSQIDGAA 94
DB 73 HORFQVALSTGAGAVASABAAAA 95

RESULT 5
US-11-052-554A-162
; Sequence 162, Application US/11052554A
; Publication No. US2005028866A1
; GENERAL INFORMATION:
; APPLICANT: Sachdeva, et al.
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
; FILE REFERENCE: 30853/40359A
; CURRENT APPLICATION NUMBER: US/11/052,554A
; CURRENT FILING DATE: 2005-02-07
; PRIOR APPLICATION NUMBER: US 60/589,227
; PRIOR FILING DATE: 2004-07-20
; PRIOR APPLICATION NUMBER: IN 173/DEL/2004
; PRIOR FILING DATE: 2004-02-06
; NUMBER OF SEQ ID NOS: 763
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 162
; LENGTH: 543
; TYPE: PRF
; ORGANISM: Mycobacterium tuberculosis H37Rv
US-11-052-554A-162

Query Match      17.8%; Score 85.5; DB 7; Length 543;
Best Local Similarity 30.7%; Pred. No. 0.37;
Matches 27; Conservative 10; Mismatches 44; Indels 7; Gaps 1;

QY 11 ADIGTQVSDNALHGVTAAGTATGTVGTPAGADEVSAQAATAFTSEGIQLLASNSAQD 70
DB 19 ANIGTSTLS-----AANAAALAAFTTGTVAAGADEVSQAIALFSDVATHTYQSLNQAAA 71

QY 71 QDHRAGEAVQDVARTYSQIDGAGVFA 98
DB 72 FHHSFVQTINAAAGAVSSAANAASQA 99

RESULT 6
US-11-052-554A-166
; Sequence 166, Application US/11052554A
; Publication No. US2005028866A1
; GENERAL INFORMATION:
; APPLICANT: Sachdeva, et al.
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
; FILE REFERENCE: 30853/40359A
; CURRENT APPLICATION NUMBER: US/11/052,554A
; CURRENT FILING DATE: 2005-02-07
; PRIOR APPLICATION NUMBER: US 60/589,227
; PRIOR FILING DATE: 2004-07-20
; PRIOR APPLICATION NUMBER: IN 173/DEL/2004
```

```
; PRIOR FILING DATE: 2004-02-06
; NUMBER OF SEQ ID NOS: 763
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 166
; LENGTH: 801
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis H37RV
US-11-052-554A-166

Query Match          17.8%; Score 85.5; DB 7; Length 801;
Best Local Similarity 37.9%; Pred. No. 0.63;
Matches 22; Conservative 10; Mismatches 19; Indels 7; Gaps 1;

Qy      11 ADIGTQVSDNALHGVTAGTALTSTVGLVPAGADEVSAQAATAFTSGTIGQLASNAS 68
Db      19 AGIGSAIS-----SANAATAAVTTGTLTGAGDEVSAAALFAGAGQAVQYASQA 69

RESULT 7
US-11-052-554A-164
; Sequence 164, Application US/11052554A
; Publication No. US20050288866A1
; GENERAL INFORMATION:
; APPLICANT: Sachdeva, et al.
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
; FILE REFERENCE: 30853/40359A
; CURRENT APPLICATION NUMBER: US/11/052,554A
; PRIOR FILING DATE: 2005-02-07
; PRIOR APPLICATION NUMBER: US 60/589,227
; PRIOR FILING DATE: 2004-07-20
; PRIOR APPLICATION NUMBER: IN 173/DEL/2004
; NUMBER OF SEQ ID NOS: 763
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 164
; LENGTH: 576
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis H37RV
US-11-052-554A-164

Query Match          17.3%; Score 83; DB 7; Length 576;
Best Local Similarity 32.2%; Pred. No. 0.74;
Matches 29; Conservative 14; Mismatches 25; Indels 22; Gaps 4;

Qy      11 ADIGTQVSDNALHGVTAGTALTSTVGLVPAGADEVSAQAATAFT--SGTIGQLASNAS 68
Db      19 ASIGSAIS-----TNAATAAAATTVTLAAGADEVSAVAALFSGHAGTYQALRTQAAA 71

Qy      69 ODOLHRGABAVQDVARTYSQIDGAGVFA 98
Db      72 FHQ-----QIVQITLTST-----AGAYFA 88

RESULT 8
US-11-052-554A-149
; Sequence 149, Application US/11052554A
; Publication No. US20050288866A1
; GENERAL INFORMATION:
; APPLICANT: Sachdeva, et al.
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
; FILE REFERENCE: 30853/40359A
; CURRENT APPLICATION NUMBER: US/11/052,554A
; PRIOR FILING DATE: 2005-02-07
; PRIOR APPLICATION NUMBER: US 60/589,227
; PRIOR FILING DATE: 2004-07-20
; PRIOR APPLICATION NUMBER: IN 173/DEL/2004
; NUMBER OF SEQ ID NOS: 763
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 149
; LENGTH: 853
```

```
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis H37RV
US-11-052-554A-149

Query Match          16.8%; Score 80.5; DB 7; Length 853;
Best Local Similarity 38.3%; Pred. No. 2.3;
Matches 23; Conservative 10; Mismatches 18; Indels 9; Gaps 2;

Qy      11 ADIGTQVSDNALHGVTAGTALTSTVGLVPAGADEVSAQAATAFTSG--IGQLASNAS 68
Db      19 AGIGSALAE-----ANAAATAAFTTALLAAGDEVSAAALFAGAGQAVQYASQA 71

RESULT 9
US-11-052-554A-160
; Sequence 160, Application US/11052554A
; Publication No. US20050288866A1
; GENERAL INFORMATION:
; APPLICANT: Sachdeva, et al.
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
; FILE REFERENCE: 30853/40359A
; CURRENT APPLICATION NUMBER: US/11/052,554A
; PRIOR FILING DATE: 2005-02-07
; PRIOR APPLICATION NUMBER: US 60/589,227
; PRIOR FILING DATE: 2004-07-20
; PRIOR APPLICATION NUMBER: IN 173/DEL/2004
; NUMBER OF SEQ ID NOS: 763
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 160
; LENGTH: 914
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis H37RV
US-11-052-554A-160

Query Match          16.8%; Score 80.5; DB 7; Length 914;
Best Local Similarity 34.5%; Pred. No. 2.5;
Matches 20; Conservative 11; Mismatches 20; Indels 7; Gaps 1;

Qy      11 ADIGTQVSDNALHGVTAGTALTSTVGLVPAGADEVSAQAATAFTSGTIGQLASNAS 68
Db      19 ADIGSTTAG-----ANAAATAAATTTSLAAGADEISAATLALFAGHGRAYQAAASA 69

RESULT 10
US-11-052-554A-137
; Sequence 137, Application US/11052554A
; Publication No. US20050288866A1
; GENERAL INFORMATION:
; APPLICANT: Sachdeva, et al.
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
; FILE REFERENCE: 30853/40359A
; CURRENT APPLICATION NUMBER: US/11/052,554A
; PRIOR FILING DATE: 2005-02-07
; PRIOR APPLICATION NUMBER: US 60/589,227
; PRIOR FILING DATE: 2004-07-20
; PRIOR APPLICATION NUMBER: IN 173/DEL/2004
; NUMBER OF SEQ ID NOS: 763
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 137
; LENGTH: 1660
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis H37RV
US-11-052-554A-137

Query Match          16.8%; Score 80.5; DB 7; Length 1660;
Best Local Similarity 37.8%; Pred. No. 5.5;
Matches 34; Conservative 13; Mismatches 26; Indels 17; Gaps 6;

Qy      11 ADIGTQVSDNALHGVTAGTALTSTVGLVPAGADEVSAQAATAFT--SGTIGQLASNAS 68
```

Db 19 ARIGSGVS-----AASAVAAMPTTEVVAAAGADEVSAGIAALFSAHAQCYOALSQAQAAA 71
Qy 69 -ODQ-LHRAGEAVQDVARTY--SQQIDGAA 94
Db 72 FHDPFVHTLTAA-----AKMTATEIANAAAA 97

RESULT 11

US-11-052-554A-154
; Sequence 154, Application US/11052554A
; Publication No. US20050288866A1
; GENERAL INFORMATION:
; APPLICANT: Sachdeva, et al.
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
; TITLE OF INVENTION: PROTEINS OF THERAPEUTIC POTENTIAL
; FILE REFERENCE: 30853/40359A
; CURRENT APPLICATION NUMBER: US/11/052,554A
; PRIOR FILING DATE: 2005-02-07
; PRIOR APPLICATION NUMBER: US 60/589,227
; PRIOR FILING DATE: 2004-07-20
; PRIOR APPLICATION NUMBER: IN 173/DEL/2004
; PRIOR FILING DATE: 2004-02-06
; NUMBER OF SEQ ID NOS: 763
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 154
; LENGTH: 767
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis H37Rv
US-11-052-554A-154

Query Match 16.7%; Score 80; DB 7; Length 767;
Best Local Similarity 30.5%; Pred. No. 2.2;

Matches 39; Conservative 12; Mismatches 41; Indels 36; Gaps 6;

Qy 4 MSHDPIADIGTGVSNALHGVTA--AGSTALTSTVGLVPAGADEVSAQAATAFTSEG- 58
Db 1 MSFVVVAPEVTAALASDLAGIGSTLAQANAAALPTAVVLAAGADEVSAALASHFGAHQ 60
Qy 59 -----IQLASNASADQLHRAG-----EAVQ-DVARTYSQ----- 88
Db 61 AYAQVSAQMSAFHQFQALTGAGAYAAAEAVVSAQSVEDLLAIAINRRFRITGRP 120
Qy 89 -IDGGAAG 95
Db 121 LIGDGANG 128

RESULT 12

US-11-052-554A-147
; Sequence 147, Application US/11052554A
; Publication No. US20050288866A1
; GENERAL INFORMATION:
; APPLICANT: Sachdeva, et al.
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
; TITLE OF INVENTION: PROTEINS OF THERAPEUTIC POTENTIAL
; FILE REFERENCE: 30853/40359A
; CURRENT APPLICATION NUMBER: US/11/052,554A
; PRIOR FILING DATE: 2005-02-07
; PRIOR APPLICATION NUMBER: US 60/589,227
; PRIOR FILING DATE: 2004-07-20
; PRIOR APPLICATION NUMBER: IN 173/DEL/2004
; PRIOR FILING DATE: 2004-02-06
; NUMBER OF SEQ ID NOS: 763
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 147
; LENGTH: 923
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis H37Rv
US-11-052-554A-147

Query Match 16.7%; Score 80; DB 7; Length 923;
Best Local Similarity 30.9%; Pred. No. 2.8;

Matches 30; Conservative 13; Mismatches 30; Indels 24; Gaps 5;

Qy 10 AADIG---TQVSDNALHGVTAAGSTALTSTVGLVPAGADEVSAQAATAFTS-----GI 59
Db 15 ASDLGGIASMISF-----ANMAAVRTTALAPPAADDEVSAALALFSSYAADYQTLVS 67
Qy 60 QLIASNASADQLHRAGE--AVQDVA----RTYSQI 89
Db 68 QVTAFFHVQFAQTLLTNAGQLYAVVDVGNGLVLTKEQQY 104

RESULT 13

US-11-052-554A-152
; Sequence 152, Application US/11052554A
; Publication No. US20050288866A1
; GENERAL INFORMATION:
; APPLICANT: Sachdeva, et al.
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
; TITLE OF INVENTION: PROTEINS OF THERAPEUTIC POTENTIAL
; FILE REFERENCE: 30853/40359A
; CURRENT APPLICATION NUMBER: US/11/052,554A
; PRIOR FILING DATE: 2005-02-07
; PRIOR APPLICATION NUMBER: US 60/589,227
; PRIOR FILING DATE: 2004-07-20
; PRIOR APPLICATION NUMBER: IN 173/DEL/2004
; PRIOR FILING DATE: 2004-02-06
; NUMBER OF SEQ ID NOS: 763
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 152
; LENGTH: 615
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis H37Rv
US-11-052-554A-152

Query Match 15.8%; Score 75.5; DB 7; Length 615;
Best Local Similarity 31.3%; Pred. No. 4.9;

Matches 26; Conservative 14; Mismatches 30; Indels 13; Gaps 4;

Qy 11 ADIGTQVSDNALHGVTAAGSTALTSTVGLVPAGADEVSAQAATAFT--TSEGIQLASNVASA 68
Db 19 AELGSAIS-----TANGAALPTVEVVAAADDEVSTQJALFGAHARSYQTLSTQAAA 71
Qy 69 QDQLR-PAGEAVQDVARTYSQID 90
Db 72 ---FHSRFVQALTTAAASYASVE 91

RESULT 14

US-11-052-554A-139
; Sequence 139, Application US/11052554A
; Publication No. US20050288866A1
; GENERAL INFORMATION:
; APPLICANT: Sachdeva, et al.
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
; TITLE OF INVENTION: PROTEINS OF THERAPEUTIC POTENTIAL
; FILE REFERENCE: 30853/40359A
; CURRENT APPLICATION NUMBER: US/11/052,554A
; PRIOR FILING DATE: 2005-02-07
; PRIOR APPLICATION NUMBER: US 60/589,227
; PRIOR FILING DATE: 2004-07-20
; PRIOR APPLICATION NUMBER: IN 173/DEL/2004
; PRIOR FILING DATE: 2004-02-06
; NUMBER OF SEQ ID NOS: 763
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 139
; LENGTH: 1306
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis H37Rv
US-11-052-554A-139

Query Match 15.8%; Score 75.5; DB 7; Length 1306;
Best Local Similarity 34.1%; Pred. No. 13;
Matches 29; Conservative 11; Mismatches 20; Indels 25; Gaps 4;


```

QY      15  TQVSDNALHG--VTAGSTALTSVTGLVPAGADE-----VSAQAATAP 54
      :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db      13  TAATDLAKIGSTITTAATAAAVAKVLPASADESVVAVALFGTHAQEYQTVSAQVAT-F 71
      :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|

QY      55  TSEGIQLASNASADOLHRAGEAV 79
      :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db      72  HDRFVQTLSSAAASS---YVAAEAV 92
      :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|

```

RESULT 15

```

US-11-052-554A-159
; Sequence 159, Application US/11052554A
; Publication No. US20050288866a1
; GENERAL INFORMATION:
; APPLICANT: Sachdeva, et al.
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
; TITLE OF INVENTION: PROTEINS OF THERAPEUTIC POTENTIAL
; FILE REFERENCE: 30853/40359A
; CURRENT APPLICATION NUMBER: US/11/052,554A
; CURRENT FILING DATE: 2005-02-07
; PRIOR APPLICATION NUMBER: US 60/589,227
; PRIOR FILING DATE: 2004-07-20
; PRIOR APPLICATION NUMBER: IN 173/DEL/2004
; PRIOR FILING DATE: 2004-02-06
; NUMBER OF SEQ ID NOS: 763
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 159
; LENGTH: 837
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis H37Rv
US-11-052-554A-159

```

```

Query Match      15.7%; Score 75; DB 7; Length 837;
Best Local Similarity 36.7%; Pred. No. 8.3;
Matches 22; Conservative 7; Mismatches 27; Indels 4; Gaps 1;

QY      10  AADIGTQVSDNALHGVTAGSTALTSVTGLVPAGADEVSAQAATAFTSEGIQLASNASAQ 69
      :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db      15  ATDLASLESSTIA---AANABAAANTTALLAAGADEVSTAVALLFGAHGQAYQALSSAQ 70
      :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|

```

Search completed: April 14, 2006, 18:42:11
 Job time : 5.19937 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using bw model

Run on: April 14, 2006, 17:03:40 ; Search time 53.857 Seconds

(without alignments)
1133.996 Million cell updates/sec

Title: US-10-620-246-90

Perfect score: 774
Sequence: 1 MRVNDPPAPGSDSARSRPAP.....RIDAISSFSKSVLTAVSAW 139

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: Geneseq_21:*
2: geneseqp1980s:*
3: geneseqp1990s:*
4: geneseqp2000s:*
5: geneseqp2001s:*
6: geneseqp2002s:*
7: geneseqp2003as:*
8: geneseqp2004s:*
9: geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	774	100.0	139	2	AAW72928
2	774	100.0	139	2	AAW72928
3	106.5	13.8	334	7	ABR86063
4	104.5	13.5	707	9	ADZ44698
5	104.5	13.5	707	9	ABR31351
6	104.5	13.5	1526	6	ABO14750
7	103.5	13.4	355	4	AAW40927
8	103.5	13.4	525	9	ABR31354
9	103.5	13.4	634	9	ABR31353
10	103.5	13.4	669	9	ABR31352
11	103.5	13.4	707	4	AAW3141
12	103.5	13.4	707	7	ADZ54941
13	103.5	13.4	707	7	ADZ54938
14	103.5	13.4	707	7	ADZ54944
15	103.5	13.4	707	7	ADZ54947
16	103.5	13.4	707	7	ADZ54935
17	103.5	13.4	707	7	ADZ54950
18	103.5	13.4	707	8	ADZ54950
19	103.5	13.4	707	8	ADZ54950
20	103.5	13.4	707	8	ADZ54950
21	103	13.3	487	6	ABU11668
22	102.5	13.2	1579	8	ADS97981
23	102	13.2	135	4	AAW13633
24	102	13.2	135	4	ABR31353

25 102 13.2 135 4 AAW26030
26 102 13.2 135 4 ABR27414
27 102 13.2 135 4 ABB18062
28 102 13.2 135 4 AAW65771
29 102 13.2 135 4 AAW53395
30 102 13.2 135 4 ABG47416
31 102 13.2 135 4 AAW01383
32 102 13.2 135 5 ABG35403
33 101.5 13.1 566 6 ADA24286
34 101 13.0 273 7 ABO79578
35 101 13.0 19938 6 ABR98398
36 99.5 12.9 198 7 ABO79582
37 99.5 12.9 199 8 ABO59459
38 99.5 12.9 212 3 AAB58403
39 99.5 12.9 212 6 ABR41630
40 99 12.8 132 8 ADQ66269
41 98.5 12.7 19938 6 ABR76678
42 97.5 12.6 345 7 ADR60212
43 97.5 12.6 345 9 ABA20959
44 97.5 12.6 403 8 ADI16240
45 97 12.5 198 8 ABO59460

ALIGNMENTS

RESULT 1
AAW72928
ID AAW72928 standard; protein; 139 AA.
XX
AC AAW72928;
XX
DT 21-JAN-1999 (first entry)
XX
DE Mycobacterium tuberculosis antigen RD1-ORF4.
XX
KW Mycobacterium tuberculosis; antigen; vaccine; immunological; immunogen;
XX
OS Mycobacterium tuberculosis.
XX
PN W09844119-A1.
XX
PD 08-OCT-1998.
XX
PF 01-APR-1998; 98WO-DK00132.
XX
PR 02-APR-1997; 97DK-00000376.
PR 18-APR-1997; 97US-0044624P.
PR 10-NOV-1997; 97DK-00001277.
PR 05-JAN-1998; 98US-0070488P.
XX
PA (STAT-) STATENS SERUM INST.
XX
PI Andersen P, Nielsen R, Rosenkrands I, Welling K, Rasmussen PB,
PI Oettinger T, Florio W;
XX
WP1: 1998-542705/46.
XX
DR N-PSDB; AAW65938.
XX
PT New isolated mycobacteria polypeptides and nucleic acids - used for
PT developing products for the diagnosis of or vaccination against
PT mycobacterial infections, particularly tuberculosis.
XX
PS Claim 1; Page 198; 163pp; English.
XX
CC The present sequence represents a Mycobacterium tuberculosis protein.
CC Products from the present invention, which describes protein fragments
CC and nucleic acid fragments derived from M. tuberculosis, can be used in
CC the detection of and prevention of mycobacterial infections. In
CC particular, the proteins and nucleic acids can be used for the diagnosis
CC of or vaccination against tuberculosis caused by M. tuberculosis, M.
CC africanum or M. bovis

XX Sequence 139 AA;
SQ
Query Match 100.0%; Score 774; DB 2; Length 139;
Best Local Similarity 100.0%; Pred. No. 6,8e-63;
Matches 139; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MRVNDPPAPGSDSARSAPALGDPDPASGWFDSGLVPSRPTCAASSAGLPPVPPTWL 60
1 MRVNDPPAPGSDSARSAPALGDPDPASGWFDSGLVPSRPTCAASSAGLPPVPPTWL 60
DB 61 NNDVTCSSGWSGCCIGPLISPSWPRVVAAGNMPGVLEPGEGIPKIGFVVLMLAGSR 120
61 NNDVTCSSGWSGCCIGPLISPSWPRVVAAGNMPGVLEPGEGIPKIGFVVLMLAGSR 120
QY 121 IDAIGSSFSSKSVLTAVSAM 139
121 IDAIGSSFSSKSVLTAVSAM 139
DB 121 IDAIGSSFSSKSVLTAVSAM 139

RESULT 2
AAV21945
ID AAV21945 standard; protein: 139 AA.
AC AAV21945;
XX
DT 06-SEP-1999 (first entry)
XX
DE Amino acid sequence of antigen RD1-ORF4.
XX
KM Immunogenic; Mycobacterium tuberculosis; Immune response; infection;
KM tuberculosis; fusion polypeptide; T-cell epitope; ESAT-6; MPT59; TB;
KM pharmaceutical; vaccination; M. africanum; M. bovis; CFP7A; CFP30A;
KM CFP7B; CFP19; CFP27; CFP30A; RD1-ORF; CFP10A; CFP16; CFP19; CFP23;
KM CFP25A; CFP30B; CFP7B.
XX
OS Mycobacterium tuberculosis.
XX
PN WO924577-A1.
XX
PD 20-MAY-1999.
XX
PF 08-OCT-1998; 98WO-DK00438.
XX
PR 10-NOV-1997; 97DK-00001277.
PR 05-JAN-1998; 98US-0070488P.
PR 01-APR-1998; 98WO-DK000132.
XX
PA (STAT-) STARENS SERUM INST.
PI
PI Andersen P, Skjot R;
XX
XX WPI; 1999-347282/29.
DR N-PSDB; AAX81045.
XX
XX
XX New immunogenic fragment of Mycobacterium tuberculosis.
PS
PS Example 2; Page 216; 265pp; English.
XX
XX The invention describes a substantially pure immunogenic polypeptide
XX fragment (1) from Mycobacterium tuberculosis that is able to evoke a
XX protective immune response against infections by mycobacteria belonging
XX to the tuberculosis complex. The invention provides a (1) fusion
XX polypeptide comprising at least one polypeptide fragment (1) and at least
XX one fusion partner; (2) a fusion polypeptide fragment comprising a T-cell
XX epitope from M. tuberculosis protein ESAT-6, or MPT59 and a second
XX different amino acid sequence from M. tuberculosis, and/or including a
XX sequence which protects the first amino acid sequence from in vivo
XX degradation or post-translational processing; (3) a nucleic acid fragment
XX that encodes the above polypeptides. The polypeptides and nucleic acid
XX are useful as pharmaceuticals, for diagnosis of and as antigens for
XX vaccination against TB caused by Mycobacterium tuberculosis, africanum or
XX bovis. The polypeptides are also useful for diagnosing ongoing or

CC previous sensitization in an animal with bacteria belonging to the
CC tuberculosis complex. The invention also describes the use of CFP7A or
CC CFP30A or a T-cell epitope of for the induction of a strong immune
CC response in a mammal; use of CFP7B, CFP19 or MPT59-ESAT6 or a T-cell
CC epitope of for diagnosis of TB in a mammal by performing a DTH type skin
CC test; use of CFP27, CFP30A, RD1-ORF2, RD1-ORF3, RD1-ORF5, MPT59-ESAT6,
CC ESAT6-MPT59, CFP10A, CFP16, CFP19, CFP23, CFP25A, CFP30B, CFP7B or a T-
CC cell epitope of for the preparation of an immunological composition; and
CC for the preparation of a subunit vaccine
XX
SQ Sequence 139 AA;
Query Match 100.0%; Score 774; DB 2; Length 139;
Best Local Similarity 100.0%; Pred. No. 6,8e-63;
Matches 139; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MRVNDPPAPGSDSARSAPALGDPDPASGWFDSGLVPSRPTCAASSAGLPPVPPTWL 60
1 MRVNDPPAPGSDSARSAPALGDPDPASGWFDSGLVPSRPTCAASSAGLPPVPPTWL 60
DB 61 NNDVTCSSGWSGCCIGPLISPSWPRVVAAGNMPGVLEPGEGIPKIGFVVLMLAGSR 120
61 NNDVTCSSGWSGCCIGPLISPSWPRVVAAGNMPGVLEPGEGIPKIGFVVLMLAGSR 120
QY 121 IDAIGSSFSSKSVLTAVSAM 139
121 IDAIGSSFSSKSVLTAVSAM 139
DB 121 IDAIGSSFSSKSVLTAVSAM 139

RESULT 3
AEB86063
ID AEB86063 standard; protein: 334 AA.
XX
XX AEB86063;
AC
XX
DT 06-OCT-2005 (first entry)
XX
XX Protein derived from DNA damage detection related human DNA, SEQ ID 20.
XX
XX diagnostic; screening; pharmaceutical; gene expression; mutation.
XX
XX Homo sapiens.
XX
XX JP2003180359-A.
XX
XX
XX 02-JUL-2003.
XX
XX 13-DEC-2001; 2001JP-00379607.
XX
XX 13-DEC-2001; 2001JP-00379607.
XX
XX (KAZU-) ZH KAZUSA DNA KENKYUSHO.
XX
XX WPI; 2003-793311/75.
DR N-PSDB; AEB86042.
XX
XX
XX Novel DNA derived from human adult brain, human tonsil, human adult
XX hippocampus and human fetal whole brain and encoding polypeptide, useful
XX as gene diagnostic agent.
PT
PT Claim 4; SEQ ID NO 20; 87pp; Japanese.
XX
XX The invention relates to a novel DNA derived from a human adult brain,
XX tonsil, hippocampus and human fetal brain and its encoded polypeptide, as
XX given in the specification. The novel DNA is useful as a diagnostic
XX agent. The encoded polypeptide is useful as a reagent for screening
XX pharmaceutical compounds which inhibit the biological activity of the
XX polypeptide. A gene containing the novel DNA is useful as a probe for
XX diagnosing the increased expression of damage, mutation or expression
XX reduction in the human DNA. This sequence represents the protein derived
XX from a novel human DNA sequence used in the screening and diagnostic
XX method of the invention. Note: This sequence is an embedded protein taken
XX from the coding DNA shown in the sequence listing of the specification.

CC p50, RelB:p50 and Real:p52 domains. Thus the specific binding

CC administering to the cell a PSF polypeptide, polynucleotide and/or a PSF

CC complex antagonist, a method of treating a subject with a condition
 CC mediated by a steroid receptor or characterized by an abnormality in a
 CC steroid receptor signal transduction pathway and/or an abnormal level of
 CC interaction between a PSF polypeptide and a steroid receptor by
 CC disrupting or promoting the interaction in cells or inhibiting or
 CC promoting the activity of a PSF complex, and a method of preventing or
 CC treating a condition mediated by a steroid receptor by administering a
 CC PSF polypeptide, complex and/or their agonist or antagonist. The methods
 CC are useful for modulating a steroid receptor or process mediated by a
 CC steroid receptor in a cell, for treating a subject or individual having a
 CC condition mediated by a steroid receptor or characterized by an
 CC abnormality in a steroid receptor signal transduction pathway and/or an
 CC abnormal level of interaction between a PSF polypeptide and a steroid
 CC receptor. The methods are useful for preventing or treating a condition
 CC mediated by a steroid receptor in a subject and for evaluating a
 CC substance for its ability to regulate the onset of labor by reacting a
 CC PSF polypeptide and a progesterone receptor and a test substance, where
 CC the PSF polypeptide and receptor bind to form a complex and comparing to
 CC a control in the absence of the substance to determine if the substance
 CC stimulates or inhibits the binding of the PSF polypeptide to the receptor
 CC and thus regulates the onset of labor. The methods are also useful for
 CC preventing and/or treating pre-term labor, for inducing labor in a
 CC subject, for preventing or treating a condition mediated by a
 CC progesterone receptor, for modulating the onset of labor, for preventing
 CC pre-term or premature labor, for reducing the risk of pre-term or
 CC premature labor, for stopping labor preparatory to Cesarean delivery or
 CC for controlling the timing of parturition in animals, such as domestic
 CC animals. This sequence represents a human PSF proline/glutamine rich
 CC polypeptide of the invention.

XX Sequence 707 AA;

Query Match 13.5%; Score 104.5; DB 9; Length 707;
 Best Local Similarity 28.0%; Pred. No. 0.7;
 Matches 33; Conservative 6; Mismatches 41; Indels 43; Gaps 4;

QY 6 PPAAGSDSAR-----SRPAPALGPPDPASG-----WPDGSLVPSRPICAA 45
 DB 99 PPPPPQSSKRVVAQGPGAPAGVSTPPASSAPATPPTSGAPPGSGPTTPTPPAVT 158
 QY 46 SSNAGLPVPVPTWLVNDVTCSCGWVSCCTGPIISPMPVWVAAGNMPTGVLEBEGT 105
 DB 159 SAPPAGAPPTTTP-----SSGVPTTTPQAGGAPPPAPPAVPGGP 196
 QY 106 -PKIG 109
 DB 197 GPKQG 201

RESULT 6

ABO14750 standard; protein: 1526 AA.

XX ABO14750;

DT 25-AUG-2003 (first entry)

XX Novel human protein #123.

XX Human; NOV, gene therapy; endocrine related disease; diabetes;
 KW metabolism-related disease; obesity; central nervous system disorder;
 KW Alzheimer's disease; Parkinson's disease; epilepsy; multiple sclerosis;
 KW schizophrenia; depression; autoimmune disorder; inflammatory disorder;
 KW psoriasis; allergy; lupus erythematosus; asthma; cancer;
 KW inflammatory bowel disease; rheumatoid arthritis; osteoarthritis;
 KW colon cancer; lung cancer; liver cancer; breast cancer; ovarian cancer;
 KW prostate cancer; brain cancer; melanoma; liver disease; liver cirrhosis;
 KW lung disease; emphysema; obstructive pulmonary disease; haemophilia;
 KW stroke; infection.

XX Homo sapiens.

XX MO2003023002-A2.

XX 20-MAR-2003.
 XX 09-SEP-2002; 2002MO-US028539.

PR 07-SEP-2001; 2001US-0318120P.
 PR 07-SEP-2001; 2001US-0318130P.
 PR 10-SEP-2001; 2001US-0318430P.
 PR 17-SEP-2001; 2001US-0322636P.
 PR 17-SEP-2001; 2001US-0322781P.
 PR 17-SEP-2001; 2001US-0322816P.
 PR 17-SEP-2001; 2001US-0322817P.
 PR 19-SEP-2001; 2001US-0323519P.
 PR 20-SEP-2001; 2001US-0323631P.
 PR 20-SEP-2001; 2001US-0323636P.
 PR 25-SEP-2001; 2001US-0324969P.
 PR 25-SEP-2001; 2001US-0325091P.
 PR 26-SEP-2001; 2001US-0324990P.
 PR 17-APR-2002; 2002US-0373212P.
 PR 06-SEP-2002; 2002US-00236177.

XX (CURA-) CURAGEN CORP.

PI Spytek KA, Paturajan M, Gorman L, Li L, Anderson DW, Zhong M;
 PI Gerlach VL, Vernet CM, Ellerman K, Berghs C, Rothenberg NB, Guo X;
 PI Shinkets RA, Leach MD, Catterton B, Kekuda R, Ji W, Miller CE;
 PI Rieger DK, Taupier RJ, Shenoy SG, Liu X, Padigaru M, Alsbrook JP;
 PI Lepley DW, Edinger SR, Burgess CR;

DR WPI; 2003-313242/30.
 DR N-PSDE; ACD19443.

PT New cytoplasmic, nuclear membrane bound or secreted polypeptides (NOVX)
 PT and polynucleotides, useful in gene therapy, e.g. for treating or
 PT preventing obesity, multiple sclerosis, allergy, cancers, hemophilia,
 PT stroke or infections.

XX Claim 1; Page 349; 586pp; English.

CC The invention describes a new isolated polypeptide (NOVX). The NOVX
 CC polypeptide, nucleic acid and antibody are useful as therapeutics,
 CC particularly in the manufacture of a medicament for treating a syndrome
 CC associated with a human disease, which includes a pathology associated
 CC with NOVX polypeptide. The DNA encoding the protein is useful in gene
 CC therapy for treating the disease or condition. In particular, the NOVX
 CC polypeptide or polynucleotide is useful for treating endocrine/
 CC metabolism-related diseases (e.g. obesity or diabetes), central nervous
 CC system disorders (e.g. Alzheimer's disease, Parkinson's disease,
 CC epilepsy, multiple sclerosis, schizophrenia or depression), autoimmune
 CC and inflammatory disorders (e.g. psoriasis, allergy, lupus erythematosus,
 CC asthma, inflammatory bowel disease, rheumatoid arthritis or
 CC osteoarthritis), cancers (e.g. colon, lung, liver, breast, ovarian,
 CC prostate or brain cancers, or melanoma), liver diseases (e.g. liver
 CC cirrhosis), lung diseases (emphysema or obstructive pulmonary disease),
 CC haemophilia, stroke, or infections (e.g. viral, bacterial or parasitic).
 CC These are also useful in developing powerful assay system for functional
 CC analysis of various human disorders, as well as in diagnostic
 CC applications, and for monitoring the effects of drugs during clinical
 CC trials. This is the amino acid sequence of a novel human NOV protein

XX Sequence 1526 AA;

Query Match 13.5%; Score 104.5; DB 6; Length 1526;
 Best Local Similarity 28.7%; Pred. No. 1.6;
 Matches 33; Conservative 8; Mismatches 31; Indels 43; Gaps 6;

QY 6 PPAAGSDSARPPALG--PDPASGWPDSGLVPSRPICAASSAGLPP-----VP 56
 DB 893 PPLPGAGIPPEPPLPGVGIPEPPLPG--AGIPPEPL-----PGAGIPPEPPLPGAGIP 945
 QY 57 PTLVNDVTCSCGWVSCCTGPIISPMPVWV-----AAGNMPTGVLEBEGTIP 106
 DB 946 P-----PPLPRVGIPPEPPLPGAGIPPEPPLPGAGIP 978

RESULT 7
 ID AAM40927 standard; protein, 355 AA.
 XX AAM40927;
 AC AAM40927;
 XX 22-OCT-2001 (first entry)
 DT 22-OCT-2001 (first entry)
 XX Human polypeptide SEQ ID NO 5858.
 DE Human; noctropic; immunosuppressant; cytostatic; gene therapy; cancer;
 KM peripheral nervous system; neuropathy; central nervous system; CNS;
 KM Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 KM amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 KM chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
 KM leukaemia.
 XX Homo sapiens.
 OS Homo sapiens.
 XX WO200153312-A1.
 PN 26-JUL-2001.
 PD 26-JUL-2001.
 XX 26-DEC-2000; 2000WO-US034263.
 PF 26-DEC-1999; 99US-00471275.
 XX 21-JAN-2000; 2000US-00488725.
 PR 25-APR-2000; 2000US-00552317.
 PR 20-JUN-2000; 2000US-00598042.
 PR 19-JUL-2000; 2000US-00620312.
 PR 03-AUG-2000; 2000US-00653450.
 PR 14-SEP-2000; 2000US-00662191.
 PR 19-OCT-2000; 2000US-00693036.
 PR 29-NOV-2000; 2000US-00727344.
 XX (HYSB-) HYSBQ INC.
 PA Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
 PI Zhou F, Goodrich R, Drmanac RT;
 XX WPI; 2001-442253/47.
 DR N-PSDB; AAI60083.
 XX Novel nucleic acids and polypeptides, useful for treating disorders such
 PT as central nervous system injuries.
 XX Example 2; SEQ ID NO 5858; 10078pp; English.
 XX The invention relates to human nucleic acids (AAI57798-AAI61369) and the
 CC encoded polypeptides (AAM38642-AAM42213) with noctropic,
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localised neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilization of the activities such as: Immune system suppression,
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukaemia and
 CC C.N.S disorders. Note: The sequence data for this patent did not form
 CC part of the printed specification
 XX Sequence 355 AA;
 SO Query Match 13.4%; Score 103.5; DB 4; Length 355;
 Best Local Similarity 28.0%; Pred. No. 0.42;
 Matches 35; Conservative 6; Mismatches 41; Indels 43; Gaps 4;

OY 6 PPAFGSDGAR-----SRPADALGPDPAPAG-----WPDGGLVPSRPTCA 45
 DB 18 PPPPPDSSKVVAAQCGPAPAGVGSAPPASSAPATPPTSGAPGSGPGTPTTPPAVT 77
 OY 46 SSSAGLPPPPVPTPLANDVYCCSGWVSCCIGPLISPSMPPRWVAAGWMPGVGLPGSGI 105
 DB 78 SAPPGAPPPPTP-----SSGVPTTTPQAGGPPPPAAVPGGPP 115
 OY 106 -PKIG 109
 DB 116 GPKQG 120
 RESULT 8
 ID AEB31354 standard; protein, 525 AA.
 AC AEB31354;
 AC AEB31354;
 DT 06-OCT-2005 (first entry)
 DT 06-OCT-2005 (first entry)
 XX Human SFPQ protein #2.
 DE Protein-associated splicing factor; SFPQ; PSF-A; labor; premature labor;
 KM parturition; cytostatic; gynecological; tocolytic.
 KM Homo sapiens.
 OS Homo sapiens.
 XX WO2005068501-A1.
 PN 28-JUL-2005.
 PD 14-JAN-2005; 2005WO-CA000042.
 PF 15-JAN-2004; 2004US-0536598P.
 PR 15-JAN-2004; 2004US-0536598P.
 XX (MOUN) MOUNT SINAI HOSPITAL.
 PA Lye S, Dong X;
 PI WPI; 2005-522812/53.
 DR GENBANK; AAM27708.
 XX Modulating a steroid receptor or steroid receptor process for treating
 PT onset of labor and cancer, comprises administering a polypyridimine tract
 PT binding protein-associated splicing factor polypeptide.
 XX Claim 22; SEQ ID NO 5; 91pp; English.
 XX The invention relates to a method of modulating a steroid receptor or a
 CC process mediated by a steroid receptor in a cell comprising administering
 CC a polypyridimine tract binding protein-associated splicing factor (PSF)
 CC polypeptide, a polynucleotide encoding a PSF polypeptide (PSF
 CC polynucleotide), an isolated complex of a PSF polypeptide and a steroid
 CC receptor (PSF complex) and/or their agonist or antagonist. The invention
 CC also relates to a method of inhibiting transactivation domains of a
 CC steroid receptor in a cell by administering to the cell a PSF
 CC polypeptide, polynucleotide, complex, and/or their agonist, a method of
 CC stimulating transactivation domains of a steroid receptor in a cell by
 CC administering to the cell a PSF polypeptide, polynucleotide and/or a PSF
 CC complex antagonist, a method of treating a subject with a condition
 CC mediated by a steroid receptor or characterized by an abnormality in a
 CC steroid receptor signal transduction pathway and/or an abnormal level of
 CC interaction between a PSF polypeptide and a steroid receptor by
 CC disrupting or promoting the interaction in cells or inhibiting or
 CC promoting the activity of a PSF complex, and a method of preventing or
 CC treating a condition mediated by a steroid receptor by administering a
 CC PSF polypeptide, complex and/or their agonist or antagonist. The methods
 CC are useful for modulating a steroid receptor or process mediated by a
 CC steroid receptor in a cell, for treating a subject or individual having a
 CC condition mediated by a steroid receptor or characterized by an
 CC abnormality in a steroid receptor signal transduction pathway and/or an
 CC abnormal level of interaction between a PSF polypeptide and a steroid

receptor. The methods are useful for preventing or treating a condition mediated by a steroid receptor in a subject and for evaluating a substance for its ability to regulate the onset of labor by reacting a PSF polypeptide and a progesterone receptor and a test substance, where the PSF polypeptide and receptor bind to form a complex and comparing to a control in the absence of the substance to determine if the substance stimulates or inhibits the binding of the PSF polypeptide to the receptor and thus regulates the onset of labor. The methods are also useful for preventing and/or treating pre-term labor, for inducing labor in a CC subject, for preventing or treating a condition mediated by a progesterone receptor, for modulating the onset of labor, for preventing pre-term or premature labor, for reducing the risk of pre-term or premature labor, for stopping labor preparatory to Cesarean delivery or for controlling the timing of parturition in animals, such as domestic CC animals. This sequence represents a human SFPQ protein used in the scope of the invention.

Sequence 525 AA;

Query Match 13.4%; Score 103.5; DB 9; Length 525;

Best Local Similarity 28.0%; Pred. No. 0.63; Mismatches 41; Indels 43; Gaps 4;

Matches 35; Conservative 6; Mismatches 41; Indels 43; Gaps 4;

QY 6 PPAAGSDSAR-----SRPALGPPPPASG-----WDSGLVPSRPICAA 45
 DB 99 PPPPPQDSKRVVAQGCPAPGVGAPAPASAPATPTPTSGAPGSGPPTPTPPAVT 158
 QY 46 SSSAGLPPVPVPTMTLNNDVTCGSGWVSCCIGPLISPSMPRVWVAAGNWPVGLPDEGI 105
 DB 159 SAPPGAPPPPTP-----SSGVPTTPPOAGSGPPPPPAVPPGPP 196
 QY 106 -PKIG 109
 DB 197 GPKQG 201

RESULT 9

AEB31353 ID AEB31353 standard; protein; 634 AA.

AC AEB31353;

DT 06-OCT-2005 (first entry)

XX Human SFPQ protein #1.

DB Protein-associated splicing factor; SFPQ; PSF-A; labor; premature labor;

KW parturition; cytostatic; gynecological; cocolytic.

XX Homo sapiens.

OS WO2005068501-A1.

PN 28-JUL-2005.

PD 14-JAN-2005; 2005MO-CA000042.

PR 15-JAN-2004; 2004US-0536598P.

XX (MOUN) MOUNT SINAI HOSPITAL.

XX Lye S, Dong X;

DR WPI; 2005-522812/53.

DR GENBANK; AAH04534.

XX Modulating a steroid receptor or steroid receptor process for treating

PT onset of labor and cancer, comprises administering a polypyridimine tract

XX binding protein-associated splicing factor polypeptide.

PS Claim 22; SEQ ID NO 4; 91PP; English.

XX The invention relates to a method of modulating a steroid receptor or a

process mediated by a steroid receptor in a cell comprising administering a polypyridimine tract binding protein-associated splicing factor (PSF) CC polypeptide, a polynucleotide encoding a PSF polypeptide (PSF CC polynucleotide), an isolated complex of a PSF polypeptide and a steroid CC receptor (PSF complex) and/or their agonist or antagonist. The invention CC also relates to a method of inhibiting transactivation domain of a CC steroid receptor in a cell by administering to the cell a PSF CC polypeptide, polynucleotide, complex, and/or their agonist, a method of CC stimulating transactivation domains of a steroid receptor in a cell by CC administering to the cell a PSF polypeptide, polynucleotide and/or a PSF CC complex antagonist, a method of treating a subject with a condition CC mediated by a steroid receptor or characterized by an abnormal level in a CC steroid receptor signal transduction pathway and/or an abnormal level of CC interaction between a PSF polypeptide and a steroid receptor by CC disrupting or promoting the interaction in cells or inhibiting or CC promoting the activity of a PSF complex, and a method of preventing or CC treating a condition mediated by a steroid receptor by administering a CC PSF polypeptide, complex and/or their agonist or antagonist. The methods CC are useful for modulating a steroid receptor or process mediated by a CC steroid receptor in a cell, for treating a subject or individual having a CC condition mediated by a steroid receptor or characterized by an CC abnormal level in a steroid receptor signal transduction pathway and/or an CC abnormal level of interaction between a PSF polypeptide and a steroid CC receptor. The methods are useful for preventing or treating a condition CC mediated by a steroid receptor in a subject and for evaluating a CC substance for its ability to regulate the onset of labor by reacting a CC PSF polypeptide and a progesterone receptor and a test substance, where CC the PSF polypeptide and receptor bind to form a complex and comparing to CC a control in the absence of the substance to determine if the substance CC stimulates or inhibits the binding of the PSF polypeptide to the receptor CC and thus regulates the onset of labor. The methods are also useful for CC preventing and/or treating pre-term labor, for inducing labor in a CC subject, for preventing or treating a condition mediated by a CC progesterone receptor, for modulating the onset of labor, for preventing CC pre-term or premature labor, for reducing the risk of pre-term or CC premature labor, for stopping labor preparatory to Cesarean delivery or CC for controlling the timing of parturition in animals, such as domestic CC animals. This sequence represents a human SFPQ protein used in the scope CC of the invention.

SO Sequence 634 AA;

Query Match 13.4%; Score 103.5; DB 9; Length 634;

Best Local Similarity 28.0%; Pred. No. 0.77; Mismatches 41; Indels 43; Gaps 4;

Matches 35; Conservative 6; Mismatches 41; Indels 43; Gaps 4;

QY 6 PPAAGSDSAR-----SRPALGPPPPASG-----WDSGLVPSRPICAA 45
 DB 26 PPPPPQDSKRVVAQGCPAPGVGAPAPASAPATPTPTSGAPGSGPPTPTPPAVT 85
 QY 46 SSSAGLPPVPVPTMTLNNDVTCGSGWVSCCIGPLISPSMPRVWVAAGNWPVGLPDEGI 105
 DB 86 SAPPGAPPPPTP-----SSGVPTTPPOAGSGPPPPPAVPPGPP 123
 QY 106 -PKIG 109
 DB 124 GPKQG 128

RESULT 10

AEB31352 ID AEB31352 standard; protein; 669 AA.

AC AEB31352;

DT 06-OCT-2005 (first entry)

XX Human PSF short form, PSF-F.

DB Protein-associated splicing factor; PSF; PSF-F; labor; premature labor;

KW parturition; cytostatic; gynecological; cocolytic.

XX Homo sapiens.

CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a human protein (shown in Table 2 of
CC the specification) which is differentially expressed during pain. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 707 AA;

Query Match 13.4%; Score 103.5; DB 7; Length 707;

Best Local Similarity 28.0%; Pred. No. 0.87; Mismatches 35; Conservative 6; Indels 43; Gaps 4;

QY 6 PPAAGSDSAR-----SRPAPALGPPDPASG-----WFDGGLVPSRPICAA 45
DB 99 PPPPPQDSSKRVVAGGPGAPGVGSAPPASAPPAATPPPTSGAPGSGPPTPPPAVT 158
QY 46 SSSAGLPPVPVPTWLNNDVTCGSGWVSCCIGPLISPMRWVWAAGNMPGVGLPBGCI 105
DB 159 SAPPGAPPPTRP-----SSGVPTTPPGAAGPPPPPAVPGPGP 196
QY 106 -PKIG 109
DB 197 GPKQG 201

RESULT 14

ADE54944 ID ADE54944 standard; protein: 707 AA.

AC ADE54944;

DT 29-JAN-2004 (first entry)

DE Human Protein P23246, SEQ ID NO 749.

XX Human; pain; neuronal tissue; gene therapy;
KW spinal segmental nerve injury; chronic constriction injury; CCI;
KM spared nerve injury; SNI; Chung.

OS Homo sapiens.

PN WO2003016475-A2.

PD 27-FEB-2003.

PR 14-AUG-2002; 2002MO-US025765.

PR 14-AUG-2001; 2001US-0312147P.

PR 01-NOV-2001; 2001US-0346382P.

PR 26-NOV-2001; 2001US-0333347P.

XX (GENO) GEN HOSPITAL CORP.
PA (FARB) BAYER AG.

PI Woolf C, D'Urso D, Belfort K, Costigan M;
XX WPI; 2003-268312/26.
DR GENBANK; P23246.

PT New composition comprising two or more isolated polypeptides, useful for
PT preparing a medicament for treating pain in an animal.

PS Claim 1; Page; 1017pp; English.

CC The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a human protein (shown in Table 2 of
CC the specification) which is differentially expressed during pain. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 707 AA;

Query Match 13.4%; Score 103.5; DB 7; Length 707;

Best Local Similarity 28.0%; Pred. No. 0.87; Mismatches 35; Conservative 6; Indels 41; Gaps 4;

QY 6 PPAAGSDSAR-----SRPAPALGPPDPASG-----WFDGGLVPSRPICAA 45
DB 99 PPPPPQDSSKRVVAGGPGAPGVGSAPPASAPPAATPPPTSGAPGSGPPTPPPAVT 158
QY 46 SSSAGLPPVPVPTWLNNDVTCGSGWVSCCIGPLISPMRWVWAAGNMPGVGLPBGCI 105
DB 159 SAPPGAPPPTRP-----SSGVPTTPPGAAGPPPPPAVPGPGP 196
QY 106 -PKIG 109
DB 197 GPKQG 201

RESULT 15

ADE54947 ID ADE54947 standard; protein: 707 AA.

AC ADE54947;

DT 29-JAN-2004 (first entry)

DE Human Protein P23246, SEQ ID NO 752.

XX Human; pain; neuronal tissue; gene therapy;
KW spinal segmental nerve injury; chronic constriction injury; CCI;
KM spared nerve injury; SNI; Chung.

OS Homo sapiens.

PN WO2003016475-A2.
 XX
 PD 27-FEB-2003.
 XX
 PF 14-AUG-2002; 2002MO-US025765.
 XX
 PR 14-AUG-2001; 2001US-0312147P.
 PR 01-NOV-2001; 2001US-0346382P.
 PR 26-NOV-2001; 2001US-0333347P.
 XX
 PA (GEHO) GEN HOSPITAL CORP.
 PA (FARB) BAYER AG.
 XX
 PI Woolf C, D'urso D, Befort K, Costigan M;
 XX
 DR WPI; 2003-268312/26.
 DR GENBANK; P23246.
 XX
 PT New composition comprising two or more isolated polypeptides, useful for
 XX preparing a medicament for treating pain in an animal.
 PS Claim 1; Page; 1017pp; English.
 CC The invention discloses a composition comprising two or more isolated rat
 CC or human polynucleotides or a polynucleotide which represents a fragment,
 CC derivative or allelic variation of the nucleic acid sequence. Also
 CC claimed are a vector comprising the novel polynucleotide, a host cell
 CC comprising the vector, a method for identifying a nucleotide sequence
 CC which is differentially regulated in an animal subjected to pain and a
 CC kit to perform the method, an array, a method for identifying an agent
 CC that increases or decreases the expression of the polynucleotide sequence
 CC that is differentially expressed in neuronal tissue of a first animal
 CC subjected to pain, a method for identifying a compound which regulates
 CC the expression of a polynucleotide sequence which is differentially
 CC expressed in an animal subjected to pain, a method for identifying a
 CC compound that regulates the activity of one or more of the
 CC polynucleotides, a method for producing a pharmaceutical composition, a
 CC method for identifying a compound or small molecule that regulates the
 CC activity in an animal of one or more of the polypeptides given in the
 CC specification, a method for identifying a compound useful in treating
 CC pain and a pharmaceutical composition comprising the one or more
 CC polypeptides or their antibodies. The polynucleotide or the compound that
 CC modulates its activity is useful for preparing a medicament for treating
 CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
 CC therapy). The sequence presented is a human protein (shown in Table 2 of
 CC the specification) which is differentially expressed during pain. Note:
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic form directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 707 AA;

Query Match 13.4%; Score 103.5; DB 7; Length 707;
 Best Local Similarity 28.0%; Pred. No. 0.87;
 Matches 35; Conservative 6; Mismatches 41; Indels 43; Gaps 4;

QY 6 PPAPGSDSAR-----SRPAPAGDPDPASG-----WPDGIVSRPICA 45
 DB 99 PPPPQDSSKRVAVGSGAPAPGASAPAPPTSGAPGSGGPTPTPPAVT 158
 QY 46 SSSAGLPPPVPTLNDVTCSSGWBSCIGPLISPMWRVWVAAGNWPVGVLPGEGI 105
 DB 159 SAPPAGAPPTPP-----SSGVPTTPPQAGGPPPPPAVPPGPG 196
 QY 106 -PKIG 109
 DB 197 GPKQG 201

Search completed: April 14, 2006, 17:18:42
 Job time : 55.857 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: April 14, 2006, 17:19:08 ; Search time 8.15065 Seconds
(without alignments)
1640.866 Million cell updates/sec

Title: US-10-620-246-90

Perfect score: 774

Sequence: 1 MRVNDPPAPGSDSARSRAP.....RIDAGSSFKSVLTAVSAW 139

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	103.5	13.4	707	2 A46302	PTB-associated spl
2	97.5	12.6	240	2 D70894	probable pra prote
3	93.5	12.1	365	2 A39481	serum response fac
4	92	11.9	419	2 G70602	hypothetical prote
5	92	11.9	980	2 S54986	regulatory protei
6	91.5	11.8	502	2 A55197	Wiskott-Aldrich sy
7	91.5	11.8	1446	1 A45344	immediate-early pr
8	91	11.8	364	2 T48188	gene NKx6.1 protei
9	90.5	11.7	380	2 S51797	vasodilator-stimul
10	89.5	11.6	924	2 F87103	initiation factor
11	89	11.5	403	2 S52796	p19L2 protein - hu
12	89	11.5	744	2 T35192	probable ABC trans
13	88.5	11.4	451	2 A41651	retinoic acid rece
14	88	11.4	1460	1 RDB47P	immediate-early pr
15	87.5	11.3	625	2 S18420	regulatory protein
16	87.5	11.3	817	2 S51342	vep10lin - yeast
17	87	11.2	533	2 S37781	retinoid X recepto
18	87	11.2	1560	2 T00080	hypothetical prote
19	86.5	11.2	446	2 A34418	H-2 region II bind
20	86.5	11.2	448	2 DA1727	retinoid X recepto
21	86.5	11.2	520	2 R44718	RXR-beta1 isoform
22	86	11.1	742	2 T04859	hypothetical prote
23	86	11.1	839	2 T04859	extensin homolog P
24	85	11.1	907	2 B96636	hypothetical prote
25	85.5	11.0	1058	2 T13286	capnucino gene pr
26	85	11.0	211	2 B89716	protein P45B6.3 (i
27	85	11.0	241	2 T22216	hypothetical prote
28	84.5	10.9	745	2 S13586	triacylglycerol 11
29	84	10.9	721	2 B70766	hypothetical prote

30	84	10.9	846	2 H70599	hypothetical prote
31	84	10.9	2357	2 A59249	class VII unconven
32	83.5	10.8	670	2 F84540	hypothetical prote
33	83	10.7	342	2 S18649	homeotic protein H
34	83	10.7	564	2 H70804	hypothetical prote
35	82.5	10.7	129	2 T06174	p2540 protein - ba
36	82.5	10.7	576	2 T36729	probable serine/th
37	82.5	10.7	838	1 VGBEG1	glycoprotein H pre
38	82.5	10.7	1255	2 T31065	diaphanous protein
39	82.5	10.7	1278	2 T27925	hypothetical prote
40	82	10.6	907	2 A45560	sporozoite surface
41	82	10.6	1213	2 A41724	limb deformity (1d
42	81.5	10.5	133	2 C86473	arabidogalactan-pr
43	81.5	10.5	319	2 F75420	hypothetical prote
44	81.5	10.5	373	2 A47234	homeobox protein H
45	81.5	10.5	1820	2 A55494	latent transformin

ALIGNMENTS

RESULT 1

A46302
PTB-associated splicing factor, long form - human
N/Alternate names: myoblast cell surface antigen 24.1D5; polypyrimidine tract-binding f
C/Species: Homo sapiens (man)
C/Date: 21-Sep-1993 #sequence revision 18-Nov-1994 #text_change 31-Dec-2004
R/Accession: A46302; A43557; S29995
R/Patton, J.G.; Porro, E.B.; Galceran, J.; Tempst, P.; Nadal-Ginard, B.
Genes Dev. 7, 393-406, 1993
A/Title: Cloning and characterization of PSF, a novel pre-mRNA splicing factor.
A/Reference number: A46302; MUID:93194059; PMID:8449401
A/Accession: A46302
A/Molecule type: mRNA
A/Residues: 1-707 <PRT>
A/Cross-references: UNIPROT:P23246; UNIPARC:UPI00001358B9; EMBL:X70944; NID:G38457; PID
A/Note: sequence extracted from NCBI backbone (NCBI:P127206)
R/Gower, H.J.; Moore, S.B.; Dickson, G.; Elsom, V.L.; Nayak, R.; Walsh, F.S.
Development 105, 723-731, 1989
A/Title: Cloning and characterization of a myoblast cell surface antigen defined by 24
A/Reference number: A43557; MUID:90091812; PMID:2480877
A/Accession: A43557
A/Status: preliminary
A/Molecule type: mRNA
A/Cross-references: UNIPARC:UPI000016A1B4; GB:X16850; NID:G23711; PIDN:CAJ34747.1; PID:
C/Genetics:
A/Gene: GDB:SPQ; PSF
A/Cross-references: GDB:138275
A/Map position: 4q-4q
C/Keywords: alternative splicing; pre-mRNA splicing; surface antigen
P1298-359/Domain: ribonucleoprotein repeat homology <RNM1>
F1372-438/Domain: ribonucleoprotein repeat homology <RNM2>

Query Match 13.4%; Score 103.5; DB 2; Length 707;
Best Local Similarity 28.0%; Pred. No. 0.35;
Matches 35; Conservative 6; Mismatches 41; Indels 43; Gaps 4;

QY	6	PPAPGDSAR-----SRPAPALGPPPPAG-----WPDGIVPSRPICAA 45
DB	99	PPPPPDSSKRVVACGPGAPVGVGAPPASAPAPAPPTGAPGSGPGTPTPPPAVT 158
QY	46	SSSAGLPVPVPTMLNNDVTCSSGMVSCCIGPLISPMRWVAAGNWPGVBLPGGI 105
DB	159	SAPPGAPPPTP-----SSGVPTTPPGAGGPPPPAAVPGGP 196
QY	106	-PKIG 109
DB	197	GPKGG 201
RESULT 2		
		D70894

Db 582 SIREKXKXKAL--HMDKVDTPQVTV 605

RESULT 6

A:55197

Wiskott-Aldrich syndrome protein WASP - human
C:Species: Homo sapiens (man)

C>Date: 23-Mar-1995 #sequence_revision 24-Nov-1999 #text_change 09-Jul-2004
C:Accession: A54747; A55197; I38931

R:Derry, U.M.; Ochse, H.D.; Francke, U.
Cell 78, 635-644, 1994

A:Title: Isolation of a novel gene mutated in Wiskott-Aldrich syndrome.

A:Reference number: A54747; MUID:94349367; PMID:8069912

A:Accession: A54747

A:Molecule type: mRNA

A:Residues: 1-328; 'A', 330-366; 'LHHHPLQLDVLHCHPLPRLVGHPCGH', 'RRHHRRRPRAPGMDQPLPSLLWC

A:Cross-references: UNIPROT:P42768; UNIPARC:UPI000017C421; GB:U12707

A>Note: this sequence is corrected in reference A55197

R:Derry, U.M.; Ochse, H.D.; Francke, U.
Cell 79, 922a, 1994

A:Reference number: A55197

A:Contents: erratum

A:Accession: A55197

A:Molecule type: mRNA

A:Residues: 1-424; 'PG', 427-502 <DER>

A:Cross-references: UNIPARC:UPI000017C422; GB:U12707; NID:G695150

A>Note: the translated sequence in Genbank entry HSU12707 (PIDN:AA62663.1) differs from

R:Khan, S.P.; Hagemann, T.L.; Radtke, B.E.; Blaese, R.M.; Rosen, F.S.
Proc. Natl. Acad. Sci. U.S.A. 92, 4706-4710, 1995

A:Title: Identification of mutations in the Wiskott-Aldrich syndrome gene and characteri

A:Reference number: I38931; MUID:95273432; PMID:7753869

A:Accession: I38931

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-502 <KMA>

A:Cross-references: UNIPARC:UPI000003CA0A; EMBL:U19927; NID:G854672; PIDN:AAC50140.1; PI

C:Genetics:

A:Gene: GDB:WAS; IMD2; WASP

A:Cross-references: GDB:120736; OMIM:301000

A:Map position: Xp11.23-Xp11.22

A>Note: detects in this gene may result in Wiskott-Aldrich syndrome

C:Keywords: Immunodeficiency

Query Match 11.8%; Score 91.5; DB 2; Length 502;
Best Local Similarity 27.9%; Pred. No. 2.5;
Matches 36; Conservative 5; Mismatches 43; Indels 45; Gaps 5;

QY 6 PPAAGSDASRRAPAL-----GPPASGMPDSGLVSPRPTCAASSAGLPP 53

Db 329 PPIVGNKGRSGPLPVLGIAPPPPTPRGPPPRGSG-----PPPPPPATGRSGPLP 383

QY 54 PVP-----PTWLNDVTCSSGWSGCCIGPLISPSMPRWVAAGWMPGVLPGEIGIKIG 109

Db 384 PPPGAGP-----PMPPPPPPPPPPSSGNGPAPPLPPLVPAAG 424

QY 110 FVVLMTAPG 118

Db 425 -----LAPG 428

RESULT 7

A:45344

Immediate-early protein - guinea herpervirus 1 (strain Kaplan)

C:Species: guinea herpervirus 1

C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-Jul-2004

C:Accession: A45344

R:Vicek, C.; Kozmik, Z.; Paces, V.; Schlim, S.; Schwyzler, M.
Virology 179, 365-377, 1990

A:Title: Pseudorabies virus immediate-early gene overlies with an oppositely oriented op

A:Reference number: A45344; MUID:91021039; PMID:2171211

A:Accession: A45344

A:Status: translation not shown

A:Molecule type: DNA

A:Residues: 1-1446 <VIC>
A:Cross-references: UNIPROT:P33479; UNIPARC:UPI000012D219; GB:M34651; NID:G334070; PIDN

C:Superfamily: herpesvirus immediate-early protein IE1/5

C:Keywords: DNA binding; early protein; transcription regulation

Query Match 11.8%; Score 91.5; DB 1; Length 1446;
Best Local Similarity 27.4%; Pred. No. 6.7;
Matches 32; Conservative 11; Mismatches 35; Indels 39; Gaps 5;

QY 2 RVNDPPAGSDASRRAPALGPPPPASGMPDSGLVSPRPTCAASSAGLPPVP----- 57

Db 413 RAGPPSP-----PAPAAPPSAS-----ASSAASAPAPDEPARP 451

QY 58 -----TWLNDVTCSSGWSGCCIGPLISPSMPRWVAAGWMPGVLPGEIGIKIG 109

Db 452 PRKRKRSTNHTLSLMDAGPPPTDGLPLP-----LGEWPF-GSDPPADGRVRYG 499

RESULT 8

A:14818

Gene NKx6.1 protein - golden hamster

C:Species: Mesocricetus auratus (golden hamster)

C>Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 05-Oct-2004

C:Accession: I4818

R:Rudnick, A.; Lang, T.Y.; Odagiri, H.; Rutter, W.J.; German, M.S.
Proc. Natl. Acad. Sci. U.S.A. 91, 12203-12207, 1994

A:Title: Pancreatic beta cells express a diverse set of homeobox genes.

A:Reference number: I48185; MUID:95083670; PMID:7991607

A:Accession: I48188

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-364 <RES>

A:Cross-references: UNIPROT:Q6054; UNIPARC:UPI00001301FA; EMBL:X81409; NID:G587466; PI

C:Genetics:

A:Gene: NKx6.1

C:Keywords: DNA binding; homeobox; nucleus; transcription regulation

F:237-293/Domain: homeobox homology <HOX>

Query Match 11.8%; Score 91; DB 2; Length 364;
Best Local Similarity 24.3%; Pred. No. 2;
Matches 46; Conservative 9; Mismatches 30; Indels 104; Gaps 10;

QY 6 PPAAGSDASRRAPALGP-----DPPASGMPDS-----GL-----VPSR 40

Db 47 PPSSSSSSSSSSPPLGANPPGLKPPAAGLSLSSPPQGLSATPRHGINILSRPSM 106

QY 41 PI-----CAASSAGL----- 51

Db 107 PVASGALPSASPSGSSSSSSSSASATSASAAAAAASPPAGLACLPFRPSL 166

QY 52 -PPVPPPTWLNDVTCSSGWSGCCIGPLISPSMPRWVAAGWMPGVLPGEIGIKIG 109

Db 167 SPPEPPP-----GLYFSPS-----AAVAAGVPRKPLAELPGRA----- 201

QY 110 FVVLMTAPG 118

Db 202 -PIFW--PG 207

RESULT 9

A:551797

Vasodilator-stimulated phosphoprotein - human

C:Species: Homo sapiens (man)

C>Date: 15-Jul-1995 #sequence_revision 09-Mar-1996 #text_change 09-Jul-2004

C:Accession: S51797; S55526

R:Haefliger, C.; Varchau, T.; Reinhard, M.; Hoppe, J.; Lohmann, S.M.; Walter, U.
EMBO J. 14, 19-27, 1995

A:Title: Molecular cloning, structural analysis and functional expression of the prolin

A:Reference number: S51796; MUID:95129547; PMID:7828592

A:Accession: S51797

A:Molecule type: mRNA

A:Residues: 1-380 <HAF1>

A:Cross-references: UNIPROT:P50552; UNIPARC:UPI000001C69; EMBL:Z46389; NID:G624963; PI

A:Accession: S55526
 A:Molecule type: protein
 A:Residues: 11-33/87-96/140-154/255-282/297-303/305-322 <HAF2>
 A:Cross-references: UNIPARC:UPI000017C415, UNIPARC:UPI000017C416, UNIPARC:UPI000017C417,
 C:Keywords: phosphoprotein

Query Match 11.7%; Score 90.5; DB 2; Length 380;
 Best Local Similarity 42.3%; Pred. No. 2.3;
 Matches 22; Conservative 1; Mismatches 16; Indels 13; Gaps 2;

QY 6 PPAFGSDARSAPALGPPPPASGWFDSGLVPSRPICAASSAGLPPVP 57
 DB 171 PPPPG-----PPPPGPPPP-----PGLPSPGVPAAHAGAGGPPAP 209

RESULT 10

Initiation factor IF-2 [imported] - Mycobacterium leprae

C/Species: Mycobacterium leprae

C/Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004

C/Accession: P87103

R./Cole, S.T.; Eiglmeyer, K.; Parkhill, J.; James, K.D.; Thomson, N.R.; Wheeler, P.R.; Ho

R.; Davies, R.M.; Devlin, K.; Dutnoy, S.; Feltwell, T.; Frazer, A.; Hamlin, N.; Holroyd,

eam, M.A.; Rutherford, K.W.

Nature 409, 1007-1011, 2001

A:Authors: Rutter, S.; Seeger, K.; Simon, S.; Simmonds, M.; Skelton, J.; Squares, R.; Sc

A>Title: Massive gene decay in the leprosy bacillus.

A:Reference number: A86909; MUID:21128732; PMID:11234002

A:Accession: P87103

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-924 <STO>

A:Cross-references: UNIPROT:Q9Z519, UNIPARC:UPI000012D2E0; GB:AL450380; NID:913093370; F

C/Genetic:

A:Gene: INF8

C:Superfamily: translation initiation factor IF-2; translation elongation factor Tu hom

Query Match

Best Local Similarity 11.6%; Score 89.5; DB 2; Length 924;
 Matches 42; Conservative 4; Mismatches 36; Indels 47; Gaps 8;

QY 7 PAFGSDARSAPALGPPPPASGWFDSGLVPSRPICAASSAGLPPVPPTWLNDVTC 66
 DB 162 PHPG-----PAKPGGRPPRVG-----NNPFSSAOSVA-RPIPRPP----- 196

QY 67 CSGWSSCCIGPLISPS--WPRVWVAAGNMP---TGVBLPGSGIPKIGFVLMAP-GS 119
 DB 197 -----APRPSASPSMSFPPGAVGGGPRPRITGVPRPGGRPG-----APVGG 241

QY 120 RIDAIGSSP 128
 DB 242 RSDAGGNT 250

RESULT 11

S52796

CDL2 protein - human (fragment)

C/Species: Homo sapiens (man)

C/Date: 19-May-1995 #sequence_revision 10-Nov-1995 #text_change 10-Sep-1997

C/Accession: S52796

R./Ruhlmann, A.; Kreidewies, S.; Nordheim, A.

submitted to the EMBL Data Library, March 1995

A:Reference number: S52796

A:Accession: S52796

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-403 <RUH>

A:Cross-references: UNIPARC:UPI000017C32C; EMBL:X86019; NID:9762950; PID:9762951

Query Match 11.5%; Score 89; DB 2; Length 403;
 Best Local Similarity 36.6%; Pred. No. 3.3;
 Matches 26; Conservative 1; Mismatches 20; Indels 24; Gaps 3;

QY 6 PPAFGSDARSAPALGPPPPASGWFDSGLVPSRPICAASSAGLPPVP--TWLN-- 62
 DB 312 PPSGVDSPSPSGRPPPLPDRP-----SAGAPPPPPSTSTRNG 351

QY 63 -DVTCCSGWVS 72
 DB 352 QDSPCEDWES 362

RESULT 12

T35192

probable ABC transporter - Streptomyces coelicolor

C/Species: Streptomyces coelicolor

C/Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 09-Jul-2004

C/Accession: T35192

R./Seeger, K.; Harris, D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.

submitted to the EMBL Data Library, April 1998

A:Reference number: Z21571

A:Accession: T35192

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-744 <SEB>

A:Cross-references: UNIPROT:Q69995; UNIPARC:UPI00000DAC12; EMBL:AL022374; PIDN:CAA18516

A:Experimental source: strain A3(2)

C/Genetic:

A:Gene: SCQEDB:SC5B8.08

Query Match 11.5%; Score 89; DB 2; Length 744;
 Best Local Similarity 32.8%; Pred. No. 5.8;
 Matches 40; Conservative 6; Mismatches 32; Indels 44; Gaps 7;

QY 7 PAF-----GSDARSAPALGPPPPASGWFDSGLVPSRPICAASSAGLPPVPPTWLNN 62
 DB 373 PAFVSGHGPAAAPSPAPAGPSPASG--PSAPAPGPAPAPAGPSPAP----- 425

QY 63 DVTCCSGWVSCTCIGPLISPSWPRVWVAAGNMPYGVBLPGSGIPKIGFVLMAPG-SRI 121
 DB 426 -----AGGGS--APAPGSEEP---ASGSPAP-----APGSPAL 453

QY 122 DA 123
 DB 454 DA 455

RESULT 13

A41651

retinoic acid receptor coregulator - rat

C/Species: Rattus norvegicus (Norway rat)

C/Date: 28-May-1992 #sequence_revision 28-May-1992 #text_change 02-Aug-2002

C/Accession: A41651

R./Yu, V.C.; Delbert, C.; Andersen, B.; Holloway, J.M.; Devary, O.V.; Naeae, A.M.; Kim,

Cell 67, 1251-1266, 1991

A>Title: RXRbeta: a coregulator that enhances binding of retinoic acid, thyroid hormone

A:Reference number: A41651; MUID:92103690; PMID:1662118

A:Accession: A41651

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-451 <YUA>

A:Cross-references: UNIPARC:UPI0000170BPE; GB:M81766; NID:9206602; PIDN:AAA42025.1; PID

C:Superfamily: retinoic acid receptor alpha; etda transforming protein homology

C:Keywords: zinc finger

F/121-367/Domain: etda transforming protein homology <ERBA>

Query Match 11.4%; Score 88.5; DB 2; Length 451;
 Best Local Similarity 27.7%; Pred. No. 4;
 Matches 36; Conservative 12; Mismatches 43; Indels 39; Gaps 6;

QY 6 PPA--PDSARSAPALGPPPPASGWFDSGLVPSRPICAASSAGLPPVPPTWLND 63
 DB 24 PPSPPGPPTTPAPAPPPMPPLG-----PPVSSWSGSGLPAPAPGPF----- 71

QY 64 VTCSSGWSCTCIGPLISPSWPRVWVAAGNMPYGVBLPGSGIPKIGFVLMAP----- 117

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: April 14, 2006, 17:04:10 ; Search time 51.2864 Seconds
(without alignments)
1912.171 Million cell updates/sec

Title: US-10-620-246-90
Perfect score: 774
Sequence: 1 MRVNDPPAPCSDSARSRAP.....RIDAIGSSFSKSVLTAVSAW 139

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : UniProt_05.80.*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	106.5	13.8	334	Q8TF16_HUMAN	Q8TF16 homo sapien
2	104.5	13.5	707	Q86VG2_HUMAN	Q86VG2 homo sapien
3	104.5	13.5	1865	Q5VUJ7_HUMAN	Q5VUJ7 homo sapien
4	104	13.4	836	Q5VNA2_NOCRA	Q5VNA2 nocardia fa
5	103.5	13.4	525	Q6PIX2_HUMAN	Q6PIX2 homo sapien
6	103.5	13.4	634	Q9BSV4_HUMAN	Q9BSV4 homo sapien
7	103.5	13.4	707	SFPQ_HUMAN	P23246 h splicing
8	103.5	13.4	707	Q5S271_HUMAN	Q5S271 homo sapien
9	101	13.0	1204	Q4S986_TETNG	Q4S986 tetradon n
10	99	12.8	132	Q6ZYX4_HUMAN	Q6ZYX4 homo sapien
11	99	12.8	494	Q4S6R8_TETNG	Q4S6R8 tetradon n
12	99	12.8	893	Q6ZCX3_ORYSA	Q6ZCX3 oryza sativ
13	98.5	12.7	396	Q9AF00_9ACTO	Q9AF00 frankia sp.
14	98.5	12.7	437	Q8VK97_MYCTU	Q8VK97 mycobacteri
15	98.5	12.7	1332	Q59GF6_HUMAN	Q59GF6 homo sapien
16	98	12.7	609	Q8FMS3_CORER	Q8FMS3 corynebacte
17	98	12.7	7192	Q5ZIN5_NOCFA	Q5ZIN5 nocardia fa
18	97.5	12.6	240	PRA_MYCTU	Q53426 mycobacteri
19	97.5	12.6	668	Q4S8K6_TETNG	Q4S8K6 tetradon n
20	97	12.5	354	Q6Z8N9_ORYSA	Q6Z8N9 oryza sativ
21	97	12.5	655	Q51HA7_MAGER	Q51HA7 magnaporthe
22	96.5	12.5	889	Q9F2N5_STRCO	Q9F2N5 streptomyce
23	96.5	12.5	1049	Q7LC09_HUMAN	Q7LC09 homo sapien
24	96.5	12.5	1064	Q94957_HUMAN	Q94957 homo sapien
25	96	12.4	341	Q86811_GAPIC	Q86811 plasmodium
26	96	12.4	796	Q5TV91_HUMAN	Q5TV91 homo sapien
27	96	12.4	902	RUSC1_HUMAN	Q5Y9V2 homo sapien
28	96	12.4	902	Q5TV92_HUMAN	Q5TV92 homo sapien
29	96	12.4	3889	Q6SSS8_CHLRE	Q6SSS8 chlamydomon
30	95.5	12.3	153	Q5TMN1_ANOGA	Q5TMN1 anopheles g
31	95.5	12.3	540	Q51EM7_ENTHI	Q51EM7 entamoeba h

32	95	12.3	373	1	ZEN_DROSU	Q24648 drosophila
33	95	12.3	958	2	Q7JXX2_DROME	Q7JXX2 drosophila
34	95	12.3	1040	2	Q81R84_DROME	Q81R84 drosophila
35	95	12.3	1300	2	Q4Q897_LEIMA	Q4Q897 leishmania
36	95	12.3	1779	2	Q53IC9_ORYSA	Q53IC9 oryza sativ
37	94.5	12.2	298	2	Q5QK11_ORYSA	Q5QK11 oryza sativ
38	94.5	12.2	595	2	Q6YX66_ORYSA	Q6YX66 oryza sativ
39	94	12.1	370	2	Q80U25_MOUSE	Q80U25 mus musculu
40	94	12.1	754	2	Q4ST14_TETNG	Q4ST14 tetradon n
41	94	12.1	1172	2	Q4S882_TETNG	Q4S882 tetradon n
42	93.5	12.1	365	1	MEP28_HUMAN	Q02080 homo sapien
43	92.5	12.0	240	2	Q7U0Q2_MYCBO	Q7U0Q2 mycobacteri
44	92.5	12.0	458	2	Q5SRM6_CRYNE	Q5SRM6 cryptococcu
45	92	11.9	419	2	Q05589_MYCTU	Q05589 mycobacteri

ALIGNMENTS

RESULT 1					
ID	Q8TF16_HUMAN	PRELIMINARY;	PRT;	334 AA.	
AC	Q8TF16	01-JUN-2002 (TREMBLrel. 21, Created)			
DT	01-JUN-2002 (TREMBLrel. 21, Last sequence update)				
DT	01-MAR-2004 (TREMBLrel. 26, Last annotation update)				
DE	KIAA1986 protein (Fragment).				
CN	Name=KIAA1986;				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae;				
OC	Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]				
RP	NUCLEOTIDE SEQUENCE.				
RC	TISSUE=Brain;				
RX	MEDLINE=21842142; PubMed=11853319;				
RA	Nagase T., Kikuno R., Ohara O.;				
RT	"Prediction of the coding sequences of unidentified human genes. XXII.				
RT	The complete sequences of 50 new cDNA clones which code for large				
RT	protein."				
RL	DNA Res. 8:319-327(2001).				
DR	EMBL; AB075866; BAB85572.1; -; mRNA.				
FT	NON TER				
SQ	SEQUENCE 334 AA; 34836 MM; 65BDED7D0A778B19 CRC64;				
Query Match 13.8%; Score 106.5; DB 2; Length 334;					
Best local similarity 30.8%; Pred. No. 1.7;					
Matches 49; Conservative 9; Mismatches 52; Indels 49; Gaps 9;					
QY	2 RVNDPP-----ARGSDARSRAPALGDPDPAAGWFGDGLVPSRPTCA-----SSSAGL 51				
DB	177 RVSEPPRAAPPAPSPPPSPPP-----PSLPAGALIDRTDGGSSBPQ 225				
QY	52 PPVPPTPLNNDVTCGWSGCCIGPLISPSMPRVVAAGGNWPTGELP-----GE 103				
DB	226 PPEPPPP-----SG-----PLPLNGLKPEFALALPPPPPGPEVKGSGGLGHEG 270				
QY	104 GIPKIGFV---VLMAPGSRIDAIGSSFS--KSVLTAVS 137				
DB	271 RGPRAVGEGPWLIVTPGKPLPVGLSFPPLKMTLAGAS 309				
RESULT 2					
ID	Q86VG2_HUMAN	PRELIMINARY;	PRT;	707 AA.	
AC	Q86VG2	01-JUN-2003 (TREMBLrel. 24, Created)			
DT	01-JUN-2003 (TREMBLrel. 24, Last sequence update)				
DT	01-MAR-2004 (TREMBLrel. 26, Last annotation update)				
DE	Splicing factor proline/glutamine rich (polypyrimidine tract binding protein associated).				
DE	Name=SFPQ;				

OC	Homo sapiens (Human) .
OC	Euryarchaeota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominiidae;
OC	Homo.
NCBI_taxid=9606;	
[1]	
RN	Nucleotide sequence.
RP	Tissue-Brain.
RC	MEDLINE=22388577; PubMed=12477932; DOI=10.1073/pnas.242603899;
EX	Strassberg R.V., Feingold B.A., Grouse L.H., Derge J.G.,
RA	Klausner R.D., Collins F.S., Wagner I., Shemen C.M., Schuler G.D.,
RA	Aleichn S.F., Zeeberg B., Burows K.H., Scheffer C.F., Bhat N.K.,
RA	Hopkins R.P., Jordan H., Moore T., Max S.I., Wang J., Heish P.,
RA	Dichtenko L., Matusina K., Farmer A.A., Rubin G.M., Hong L.,
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA	Brownstein M.J., Ueda T.B., Toshiyuki S., Carninci P., Franke C.,
RA	Rata S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA	Boesq S.A., McEwen P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huylk S.W.,
RA	Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA	Falley J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA	Butterfield V.S.N., Krzyzinski M.I., Skalska U., Smalhus D.B.,
RA	Schmerer A., Schain J.E., Jones S.J.M., Marra M.A.;
RT	"Generation and initial analysis of more than 15,000 full-length human
RL	and mouse cDNA sequences."
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
[2]	
RN	Nucleotide sequence.
RP	Tissue-Brain.
RG	NIH MGC Project;
RL	Submitted (Apr-2003) to the EMBL/GenBank/DBJ databases.
DR	HSSP; BC051192; AAH51192.1; -; mRNA.
DR	HSSP; 008583; INO8.
DR	Ensembl; ENSG00000116560; Homo sapiens.
DR	GO; GO:0003676; P:nucleic acid binding; IEA.
DR	GO; GO:0000166; P:nucleotide binding; IEA.
DR	GO; GO:0000399; P:nuclear RNA splicing, via spliceosome; IEA.
DR	InterPro; IPRO12677; a_b_plat_nuc_bd.
DR	InterPro; IPRO12975; NOPS.
DR	InterPro; IPRO05054; RNPI_RNA_bd.
DR	Pfam; PF08075; NOPS; 1.
DR	Pfam; PF00076; RRM_1; 2.
DR	SMART; SM00360; RRM; 2.
DR	PROSITE; PS0102; RRM; 2.
KW	Activator; DNA damage; DNA repair; DNA-binding; Nuclear protein;
KW	Transcription; Transcription regulation; mRNA processing;
KW	mRNA splicing.
SO	SEQUENCE 707 AA; 76188 MW; B5C287EF6FD1811 CRC64;
Query Match	13.5%; Score 104.5; DB 2; Length 707;
Best Local Similarity	28.0%; Pred. No. 5.1;
Matches	35; Conservative 6; Mismatches 41; Indels 43; Gaps 4;
OY	PPAPGSDSAR-----SRPAPALGFDPSPASG-----WPDSGLVPSRPTCA 45
DB	99 PPPPPQDSSKPVVAQGPGAPGVGSTPPASSSAPATPTSGAPPGSGPPTPPPAVT 158
OY	46 SSASGLPPVPVPTWLNNDVTCSSGNVSCIGLIPSPWRVRVAAAGMMPTGVELRGSI 105
DB	159 SAPPAAPPPTTP-----SSGVFTTTPQAAGPPPAPPAAVGPQP 196
OY	106 -PKTG 109
DB	197 GPKG 201
RESULT 3	
OSVU37_HUMAN	
OSVU37_HUMAN PRELIMINARY;	PRT; 1865 AA.
OSVU37;	

```

DN 01-FEB-2005 (TREMblrel. 29, Created)
DT 01-FEB-2005 (TREMblrel. 29, Last sequence update)
DT 13-SEP-2005 (TREMblrel. 31, Last annotation update)
DE Formin 2.
GN Name=FMN2; ORFNames=RP11-90L13.1-001;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniota; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Martin S.;
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Hall R.;
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RA Sycamore N.;
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
RN [4]
RP NUCLEOTIDE SEQUENCE.
RA Wallis J.;
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL590490; CAH70931.1; -; Genomic DNA.
DR EMBL; AL513342; CA117121.1; -; Genomic DNA.
DR EMBL; AL646016; CA117009.1; -; Genomic DNA.
DR EMBL; AL359918; CA115795.1; -; Genomic DNA.
DR EMBL; AL359918; CAH70931.1; JOINED; Genomic DNA.
DR EMBL; AL513342; CAH70931.1; JOINED; Genomic DNA.
DR EMBL; AL359918; CAH70931.1; JOINED; Genomic DNA.
DR EMBL; AL646016; CAH70931.1; JOINED; Genomic DNA.
DR EMBL; AL590490; CA115795.1; JOINED; Genomic DNA.
DR EMBL; AL513342; CA117009.1; JOINED; Genomic DNA.
DR EMBL; AL590490; CA117121.1; JOINED; Genomic DNA.
DR EMBL; AL646016; CA115795.1; JOINED; Genomic DNA.
DR EMBL; AL513342; CA115795.1; JOINED; Genomic DNA.
DR EMBL; AL359918; CA117009.1; JOINED; Genomic DNA.
DR EMBL; AL646016; CA117121.1; JOINED; Genomic DNA.
DR GO; GO:0003779; F-actin binding; IEA.
DR GO; GO:0016043; P-cell organization and biogenesis; IEA.
DR GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR InterPro; IPR000591; DEP.
DR InterPro; IPR009408; Df_FH1.
DR InterPro; IPR003104; FH2_actin_bd.
DR InterPro; IPR001220; Lectin_legb.
DR Pfam; PF06346; Df_FH1; 3.
DR Pfam; PF02181; FH2; 1.
DR SMART; SMO0498; FH2; 1.
DR PROSITE; PS50186; DEP; 1.
DR PROSITE; PS00307; LECTIN_LEGUEM_BETA; UNKNOWN; 1.
SQ SEQUENCE 1865 AA; 194918 MW; 202D47DABE981F11 CRC64;

Query Match 13.5%; Score 104.5; DB 2; Length 1865;
Best Local Similarity 28.7%; Pred. No. 13;
Matches 33; Conservative 8; Mismatches 31; Indels 43; Gaps 6;

QY 6 PPARGSASVSRPALG--PDPASQWFGSLVPSRICAASSAGLPP-----VP 56
DB 1232 PPLRGAGIPPPPLPGVGIRPPPLPG--AGIRPPPL-----PGAIRPPPLPGAGIR 1284
QY 57 PTWLNINDVTCSCGVNCCIGPLSPSPRWV-----AAGNWFVGLRFGGIR 106
DB 1285 P-----PPLPRVGIRPPPLPGAGIRPPPLPGAGIR 1317

RESULT 4
OSYNA2 NOCPA
ID OSYNA2 NOCPA PRELIMINARY; PRT; 836 AA.
AC OSYNA2;
```

```

DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein.
DE OrderedLocustNames=nfa54870;
OS Nocardia farcinica.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Nocardiaceae; Nocardia.
OX NCBI_TaxID=37329;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=IFW 10152;
RX PubMed=15466710; DOI=10.1073/pnas.0406410101;
RA Shibata T., Hattori M.;
RA Shibata T., Hattori M.;
RT "The complete genomic sequence of Nocardia farcinica IFW 10152."
RT Proc. Natl. Acad. Sci. U.S.A. 101:14925-14930(2004).
DR EMBL; AP006618; BAD60339.1; Genomic_DNA.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 836 AA; 81926 MW; ACDP5523F8651CDD CRC64;

Query Match 13.4%; Score 104; DB 2; Length 836;
Best Local Similarity 23.2%; Pred. No. 6.6;
Matches 48; Conservative 15; Mismatches 64; Indels 80; Gaps 10;

QY 3 VNDEPAPGSDSARSR-----PAP----- 20
DB 68 VGDEPPGSPPTAGRDASTGPAPVPAPVAHPPTPTVPTGTSTPGSSSAVPEGS 127
QY 21 -----ALGPD-----PPASGWFDSGLVPSRPICASSAGLPP-----VPP- 57
DB 128 PSTGSVDAAAMPDPDSBDGTASPSGWAEGAVPPRD-----AAKRGVPPSGRAEGAVPEPR 183
QY 58 TWLANDVTCSSGWSGCCIGPLISPS--W---PRVWVAAGMWPVGLRGEGIPITGVV 112
DB 184 EAMKPGAVPRPGMGVPPRAVPVPGAGMWSVPPPGW-AKPGAVPRAPASGAGAPRGTA 242
QY 113 LML-----APGSRIDAIGSSSFSKSVLT 134
DB 243 AMMKAGAGAPDPVPSAGAGRAVAT 269

RESULT 5
Q6PIX2_HUMAN PRELIMINARY; PRT; 525 AA.
ID Q6PIX2_HUMAN PRELIMINARY; PRT; 525 AA.
AC Q6PIX2;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE SFPO protein (Fragment).
GN Name=SFPO;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Placenta;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strauberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shemen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Bueltow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.B.,
RA Brownstein M.J., Udell T.B., Tothiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loguallano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McGowan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huliy S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,

```

```

RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Gilmwood J., Schmutz J., Myers R.M.,
RA Rodriguez Y.S.N., Krzywinski M.I., Skalska U., Smallus D.B.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Placenta;
RG NIH MGC Project;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC027708; AAH27708.1; mRNA.
DR Ensembl; ENSG00000116560; Homo sapiens.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR InterPro; IPR012677; a_b_plat_nuc_bd.
DR InterPro; IPR012975; NOPS.
DR InterPro; IPR00504; RNP1_RNA_bd.
DR Pfam; PF08075; NOPS; 1.
DR Pfam; PF00076; RRM_1; 2.
DR SMART; SM00360; RRM; 2.
DR PROSITE; PS0102; RRM; 2.
DR DNA damage; DNA repair.
KW NON TER 525
SQ SEQUENCE 525 AA; 55469 MW; P8AD883A531C1781 CRC64;

Query Match 13.4%; Score 103.5; DB 2; Length 525;
Best Local Similarity 28.0%; Pred. No. 4.6;
Matches 35; Conservative 6; Mismatches 41; Indels 43; Gaps 4;

QY 6 PPAFGSDSAR-----SRPAPALGPPDPASG-----WPDGGLVPSRPICAA 45
DB 99 PPPPPQDSKRVVAQGPAPGAPGASAPASSAPATPTPGAGSGGPGTPTPPPAVT 158
QY 46 SSSAGLPPPPPTPLNDVTCSSGWSGCCIGPLISPSKPRVWVAAGMWPVGLRGEGI 105
DB 159 SAPPAPPPPTP-----SSGVPTTPQAGPPPPPAVPGGP 196
QY 106 -PKIG 109
DB 197 GPKOG 201

RESULT 6
Q9BSV4_HUMAN PRELIMINARY; PRT; 634 AA.
ID Q9BSV4_HUMAN PRELIMINARY; PRT; 634 AA.
AC Q9BSV4;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE SFPO protein (Fragment).
GN Name=SFPO;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Placenta;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strauberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shemen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Bueltow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.B.,
RA Brownstein M.J., Udell T.B., Tothiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loguallano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McGowan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huliy S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,

```

RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywicki M.I., Skalek U., Smallue D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 [2]
 RN NUCLEOTIDE SEQUENCE.
 RP TISSUE=Placenta;
 RC NIH MEC Project;
 RL Submitted (MAR-2001) to the EMBL/GenBank/DDBJ databases.
 CC -1- INTERACTION:
 CC 09NC04:EXOSCS; NBExp=1; IntAct=EBI-372151, EBI-371876;
 DR EMBL; BC004534; AAH04534.2; -; mRNA.
 DR HSSP; O08583; 1N08.
 DR Intact; Q9BSV4; -;
 DR Ensembl; ENSG00000116560; Homo sapiens.
 DR GO; GO:0003676; F:nucleic acid binding; IEA.
 DR GO; GO:0000166; F:nucleotide binding; IEA.
 DR InterPro; IPR012677; a.b.plat_nuc_bd.
 DR InterPro; IPR012975; NOPS.
 DR InterPro; IPR00504; KNPI_RNA_bd.
 DR Pfam; PF08075; NOPS; 1.
 DR Pfam; PF00076; RRM_1; 2.
 DR SMART; SM00360; RRM; 2.
 DR PROSITE; PS50102; RRM; 2.
 DR DNA damage; DNA repair.
 FT NON_TER
 FT SEQUENCE 634 AA; 68631 MW; ECD7EDABE86FA00 CRC64;
 SQ
 Query Match 13.4%; Score 103.5; DB 2; Length 634;
 Best local Similarity 28.0%; Pred. No. 5.5;
 Matches 35; Conservative 6; Mismatches 41; Indels 43; Gaps 4;
 QY 6 PPAPGSDSAR-----SRPAPALGPPDPASG-----WFDGLVSPRPICAA 45
 Db PPPPPQDSKRVVAQGPPAPGVGASAPPAATPTPGAPGSGPPTPPPAVT 85
 QY 46 SSASAGLPVPPTMTNDVTCSSGWCCTGPIISPMPRVVAAGNMPGVTLRBEGI 105
 Db 86 SAAPGAPPTPTP-----SSGVPTTPQAGGPPPPAAVPPGPGP 123
 QY 106 -PKIG 109
 Db 124 GPKQG 128
 RESULT 7
 SFPQ_HUMAN STANDARD; PRT; 707 AA.
 AC P23246; P30808; Rel. 20, Created)
 DT 01-NOV-1991 (Rel. 20, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 13-SEP-2005 (Rel. 48, Last annotation update)
 DE Splicing factor, proline- and glutamine-rich (POLY(pyrimidine tract-
 DE binding protein)-associated glutamine factor) (PTB-associated splicing
 DE factor) (PSF) (DNA-binding p52/p100 complex, 100 kDa subunit) (100-kDa
 DE DNA-binding protein) (hPOMP100).
 GN Name=SPFQ; Synonym=PSF;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Eumarchontia; Primates; Catarrhini; Homnidae;
 OC Homo.
 OX NCBI_taxID=9606;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [mRNA], PARTIAL PROTEIN SEQUENCE, ALTERNATIVE
 RP SPLICING, AND FUNCTION.
 RC TISSUE=Fetal brain;
 RX MEDLINE=9319405; PubMed=8449401;
 RX Patton J.G., Porro E.B., Galceran J., Tempet P., Nadal-Ginard B.,
 RT "Cloning and characterization of PSF, a novel pre-mRNA splicing

RT factor.";
 RL Genes Dev. 7:393-406(1993).
 RN [2]
 RP PROTEIN SEQUENCE OF 48-68 AND 213-246, BLOCKAGE OF THE N-TERMINUS,
 RP DNA-BINDING, AND SUBUNIT.
 RX MEDLINE=93176127; PubMed=8439294;
 RX Zhang W.-W., Zhang L.-X., Busch R.K., Farres J., Busch H.,
 RA "Purification and characterization of a DNA-binding heterodimer of 52
 RT and 100 kDa from HeLa cells.";
 RL Biochem. J. 290:267-272(1993).
 RN [3]
 RP PROTEIN SEQUENCE OF 292-311; 415-421 AND 503-510, AND IDENTIFICATION
 RP IN US/4/6 SNRNP COMPLEXES.
 RX MEDLINE=98072280; PubMed=9409622;
 RA Tsigelkamp S., Mundt C., Achsel T., Wall C.L., Luehmann R.,
 RT "The human US snRNP-specific 100-kD protein is an RS domain-
 RT containing, putative RNA helicase with significant homology to the
 RT yeast splicing factor Prp28p.";
 RL RNA 3:1313-1326(1997).
 RN [4]
 RP NUCLEOTIDE SEQUENCE [mRNA] OF 312-707.
 RC TISSUE=Fetal skeletal muscle;
 RX MEDLINE=9091812; PubMed=2480877;
 RA Cowser H.J., Moore S.B., Dickson G., Elsom V.L., Nayak R., Walsh F.S.,
 RT "Cloning and characterization of a myoblast cell surface antigen
 RT defined by 24.1D5 monoclonal antibody.";
 RL Development 105:723-731(1989).
 RN [5]
 RP PROTEIN SEQUENCE OF 414-421 AND 427-448, SUBCELLULAR LOCATION,
 RP INTERACTION WITH SNRPA, AND IDENTIFICATION IN A SNRNP-FREE COMPLEX
 RP WITH SNRPA.
 RX MEDLINE=99063394; PubMed=9848648; DOI=10.1017/S135838298981183;
 RA Lutz C.S., Cooke C., O'Connor J.P., Kobayashi R., Alwine J.C.,
 RT "The snRNP-free U1A (SF-A) complex(es): identification of the largest
 RT subunit as PSF, the polypyrimidine-tract binding protein-associated
 RT splicing factor.";
 RL RNA 4:1493-1499(1998).
 RN [6]
 RP PROTEIN SEQUENCE OF 600-606 AND 667-677, INTERACTION WITH PTB1, AND
 RP SUBCELLULAR LOCATION.
 RX PubMed=10653975;
 RX DOI=10.1002/(SICI)1097-4644(20000315)76:4<559::AID-JCBA>3.0.CO;2-U;
 RA Weisner M., Dechac T., Gerner C., Grimm R., Potner R., Suetmann G.,
 RT "Differential nuclear localization and nuclear matrix association of
 RT the splicing factors PSF and PTB.";
 RL J. Cell. Biochem. 76:559-566(2000).
 RN [7]
 RP FUNCTION, INTERACTION WITH PRE-mRNA, AND IDENTIFICATION IN SPLICOSOME
 RP COMPLEX.
 RX PubMed=8045264;
 RA Gozani O., Patton J.G., Reed R.,
 RT "A novel set of spliceosome-associated proteins and the essential
 RT splicing factor PSF bind stably to pre-mRNA prior to catalytic step II
 RT of the splicing reaction.";
 RL EMBO J. 13:3356-3367(1994).
 RN [8]
 RP CHROMOSOMAL TRANSLOCATION WITH TFE3.
 RX PubMed=993982; DOI=10.1038/9j.onc.1201394;
 RA Clark J., Lu Y.-J., Sidhar S.K., Parker C., Gill S., Smedley D.,
 RA Hamoudi R., Linhan W.M., Shipley J., Cooper C.S.,
 RT "Fusion of splicing factor genes PSF and Nono (p54nrb) to the TFE3
 RT gene in papillary renal cell carcinoma.";
 RL Oncogene 15:2233-2239(1997).
 RN [9]
 RP INTERACTION WITH TOP1, AND IDENTIFICATION IN A COMPLEX WITH NONO AND
 RP TOP1.
 RX MEDLINE=98434520; PubMed=9756848; DOI=10.1074/jbc.273.41.26261;
 RA Straub T., Grupe P., Unse A., Lisby M., Knudsen B.R., Tange T.O.,
 RA Westergaard O., Boege F.,
 RT "The RNA-splicing factor PSF/p54 controls DNA-topoisomerase I activity
 RT by a direct interaction.";
 RL J. Biol. Chem. 273:26261-26264(1998).
 RN [10]

RP FUNCTION IN DNA UNWINDING.
 RX PubMed=10858305; DOI=10.1021/b1992898e;
 RA Straub T., Knudsen B.R., Boege F.;
 RT "SPF/p54 (nrb) stimulates 'jumping' of DNA topoisomerase I between
 RT separate DNA helices.";
 RL Biochemistry 39:7552-7558(2000).
 RN [11]
 RP FUNCTION IN TRANSCRIPTION REGULATION, AND IDENTIFICATION BY MASS
 RP SPECTROMETRY.
 RX MEDLINE=20304193; PubMed=10847580; DOI=10.1210/me.14.6.774;
 RA Urban R.J., Bodenbury Y., Kurosky A., Wood T.G., Gasic S.;
 RT "Polyuridylic tract-binding protein-associated with the polyuridylic tract
 RT negative regulator of transcriptional activity of the porcine p450bpc
 RT insulin-like growth factor response element.";
 RL Mol. Endocrinol. 14:774-782(2000).
 RN [12]
 RP FUNCTION IN HOMOLOGOUS DNA PAIRING, AND PHOSPHORYLATION.
 RX MEDLINE=20392247; PubMed=10931916; DOI=10.1093/nar/28.16.3022;
 RA Akhmedov A.T., Lopez B.S.;
 RT "Human A-to-Kaa homologous DNA-pairing protein is the splicing factor
 RT PSF and promotes DNA strand invasion.";
 RL Nucleic Acids Res. 28:3022-3030(2000).
 RN [13]
 RP FUNCTION IN NUCLEAR RETENTION OF A-TO-I EDITED RNAs, AND
 RP IDENTIFICATION IN A COMPLEX WITH NONO AND MATR3.
 RX MEDLINE=21417141; PubMed=11525732; DOI=10.1016/S0092-8674(01)00466-4;
 RA Zhang Z., Carmichael G.G.;
 RT "The fate of dsRNA in the nucleus: a p54 (nrb)-containing complex
 RT mediates the nuclear retention of promiscuously A-to-I edited RNAs.";
 RL Cell 106:465-475(2001).
 RN [14]
 RP INTERACTION WITH SNRP70, AND PHOSPHORYLATION.
 RX PubMed=11514619;
 RA Shav-Tal Y., Cohen M., Laper S., Dye B., Paton J.G.,
 RT "Vanderkerckhove J., Zlotof D.;
 RT "Nuclear relocalization of the pre-mRNA splicing factor PSF during
 RT apoptosis involves hyperphosphorylation, masking of antigenic
 RT epitopes, and changes in protein interactions.";
 RL Mol. Biol. Cell 12:2328-2340(2001).
 RN [15]
 RP FUNCTION IN TRANSCRIPTION REGULATION, AND INTERACTIONS WITH RXRA; THRA
 RP AND SIN3A.
 RX PubMed=11259580; DOI=10.1128/MCB.21.7.2298-2311.2001;
 RA Mathur M., Tucker P.W., Samuels H.H.;
 RT "PSF is a novel corepressor that mediates its effect through Sin3A and
 RT the DNA binding domain of nuclear hormone receptors.";
 RL Mol. Cell Biol. 21:2298-2311(2001).
 RN [16]
 RP FUNCTION IN TRANSCRIPTION REGULATION, INTERACTIONS WITH NR5A1 AND
 RP SIN3A, AND IDENTIFICATION IN A COMPLEX WITH NONO AND NR5A1.
 RX MEDLINE=21895433; PubMed=11897684; DOI=10.1210/en.143.4.1280;
 RA Sewer M.B., Nguyen V.O., Huang C.J., Tucker P.W., Kagawa N.,
 RT "Waterman M.R.;
 RT "Transcriptional activation of human CYP17 in H295R adrenocortical
 RT cells depends on complex formation among p54(nrb)/Nono, protein-
 RT associated splicing factor, and SF-1, a complex that also participates
 RT in repression of transcription.";
 RL Endocrinology 143:1280-1290(2002).
 RN [17]
 RP INTERACTIONS WITH NONO AND US snRNA, AND IDENTIFICATION IN IN US/4/6
 RP SNRNP AND SPLICOSOME COMPLEXES.
 RX MEDLINE=22296370; DOI=10.1017/S1358382020202070;
 RA Peng R., Dye B.T., Perez I., Barnard D.C., Thompson A.B., Paton J.G.;
 RT "PSF and p54nrb bind a conserved stem in US snRNA.";
 RL RNA 8:1334-1347(2002).
 RN [18]
 RP FUNCTION IN DNA REPAIR, IDENTIFICATION BY MASS-SPECTROMETRY,
 RP DNA-BINDING, AND SUBUNIT.
 RX PubMed=1559677; DOI=10.1074/jbc.M412758200;
 RA Bladen C.L., Udayakumar D., Takeda Y., Dymov W.S.;
 RT "Identification of the polypyrimidine tract binding protein-associated
 RT splicing factor p54(nrb) complex as a candidate DNA double-strand
 RT break rejoining factor.";

RL J. Biol. Chem. 280:5205-5210(2005).
 CC -1- FUNCTION: DNA- and RNA binding protein, involved in several
 CC nuclear processes. Essential pre-mRNA splicing factor required
 CC early in spliceosome formation and for splicing catalytic step II,
 CC probably as a heteromer with NONO. Binds to pre-mRNA in
 CC spliceosome C complex, and specifically binds to intronic
 CC polypyrimidine tracts. Interacts with US snRNA, probably by
 CC binding to a purine-rich sequence located on the 3' side of US
 CC snRNA stem 1b. May be involved in a pre-mRNA coupled splicing and
 CC polyadenylation process as component of a snRNP-free complex with
 CC snRNP/U1A. The SFPO-NONO heteromer associated with MATR3 may play
 CC a role in nuclear retention of defective RNAs. SFPO may be
 CC involved in homologous DNA pairing; in vitro, promotes the
 CC invasion of ssDNA between a duplex DNA and produces a D-100p
 CC formation. The SFPO-NONO heteromer may be involved in DNA
 CC unwinding by modulating the function of topoisomerase I/TOPI; in
 CC vitro, stimulates dissociation of TOPI from DNA after cleavage and
 CC enhances its jumping between separate DNA helices. The SFPO-NONO
 CC heteromer may be involved in DNA nonhomologous end joining (NHEJ)
 CC required for double-strand break repair and V(D)J recombination
 CC and may stabilize paired DNA ends; in vitro, the complex strongly
 CC stimulates DNA end joining, binds directly to the DNA substrates
 CC and cooperates with the Ku70/G22P1-Ku80/KRCC5 (Ku) dimer to
 CC establish a functional preligation complex. SFPO is involved in
 CC transcriptional regulation. Transcriptional repression is probably
 CC mediated by an interaction of SFPO with SIN3A and subsequent
 CC recruitment of histone deacetylases (HDACs). The SFPO-NONO/SF-1
 CC complex binds to the CYP17 promoter and regulates basal and CAMP-
 CC dependent transcriptional activity. SFPO isoform long binds to the
 CC DNA binding domain (DBD) of nuclear hormone receptors, like RXRA
 CC and probably THRA, and acts as transcriptional corepressor in
 CC absence of hormone ligands. Binds the DNA sequence 5'-CTGAGTC-3'

Query Match 13.4%; Score 103.5; DB 1; Length 707;
 Best Local Similarity 28.0%; Pred. No. 6.1;

Matches 35; Conservative 6; Mismatches 41; Indels 43; Gaps 4;

Qy	6	PPAPGDSAR-----SRPAPALGPPDPASG-----WFGSLVPSRPICAA	45
Db	99	PPPPPODSKRVVAGQGPAPGVGASAPPASSAPAPPTSGAPGSGPPTPPPAVT	158
Qy	46	SSSAGLPPVPPTPTLNDVTCGSGVCCIGPLISBPVRYVAAAGNMPVGLPREGI	105
Db	159	SAPGAPGPPPTP-----SSGVPTTPPOAGPPPPPAVPPGPGP	196
Qy	106	-PKTG 109	
Db	197	GPKG 201	

RESULT 8

OS5271_HUMAN PRELIMINARY; PRT; 707 AA.
 ID OS5271;
 AC OS5271;
 DT 01-FEB-2005 (TRMBLrel. 29, Created)
 DT 01-FEB-2005 (TRMBLrel. 29, Last sequence update)
 DT 01-FEB-2005 (TRMBLrel. 29, Last annotation update)
 DE Splicing factor proline/glutamine rich (Polypyrimidine tract binding
 DE protein associated).
 GN Name=SFPO; ORFNames=RP11-181E22.2-001;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Euteleia; Euarchontoglires; Primates; Catarrhini; Hominidae;
 OC Homo.
 NCBI_taxid=9606;
 RX NCB1 [1]
 RN NUCLEOTIDE SEQUENCE.
 RA Howden P.;
 RT Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL590434; CAIL12467.1; -, Genomic DNA.
 DR EMBL; ENSG00000116560; Homo sapiens.
 DR GO; GO:0003676; P:nucleic acid binding; IEA.
 DR InterPro; IPR012677; a_b_plait_nuc_bd.

```

DR InterPro: IPR012975; NOPS.
DR InterPro: IPR00504; RNPI_RNA_bd.
DR Pfam: PF08075; NOPS; 1.
DR Pfam: PF00076; RRM_1; 2.
DR SMART: SM00360; RRM; 2.
DR PROSITE: PSS0102; RRM; 2.
KW Activator; DNA damage; DNA repair; DNA-binding; Nuclear protein;
KW Transcription; transcription regulation; mRNA processing;
KW mRNA splicing.
SQ SEQUENCE 707 AA; 76149 MW; 6D8DEA95E235847 CRC64;

Query Match 13.4%; Score 103.5; DB 2; Length 707;
Best Local Similarity 28.0%; Pred. No. 6.1;
Matches 35; Conservative 6; Mismatches 41; Indels 43; Gaps 4;

QY 6 PPAPGSDSAR-----SRPAPALGPPPPASG-----WFDGGLVSPRPICAA 45
DB 99 PPPPPQSSSKRVVAGCGPAPGVGASAPPASAPPAATPTTSGAPGSGPTPTPPPAVT 158
DB 159 SAPGAPPPPTP-----SSGVPTTPPGAGPPPPPAVPGGP 196

QY 46 SSSAGLPVPVPTLVNDVTCGSGWVSCIGPLISPMRVWVAAGNPTGVLPGEIGI 105
DB 159 SAPGAPPPPTP-----SSGVPTTPPGAGPPPPPAVPGGP 196

QY 106 -PKTG 109
DB 197 GPKQG 201

RESULT 9
Q4S86_TETNG
ID Q4S86_TETNG PRELIMINARY; PRT; 1204 AA.
AC Q4S86;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Chromosome 3 SCAPI4700, whole genome shotgun sequence.
DS (fragment).
GN ORFNames=GSTENG0022009001;
OC Tetradon nigrorivridis (Green puffer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorphi; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Tetraodon.
OX NCBI_TaxID=99883;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Stallion O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,
RA Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
RA Micaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segutens B.,
RA Daellwa C., Salenoubat M., Levy M., Boudet N., Castellano S.,
RA Authouard V., Jubin C., Castelli V., Katinka M., Vacherie B.,
RA Blemont C., Skalli Z., Cartolico L., Poulain J., De Bernardis V.,
RA Crenaud C., Duprat S., Broctier P., Coutanceau J.P., Gouzy J.,
RA Parra G., Lardier S., Chapple C., McKernan K.J., McEwan P., Bosak S.,
RA Kellis M., Volff J.N., Guigo R., Zody M.C., Mesirov J.,
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA Lander V., Schachter V., Queller F., Saurin W., Scarpelli C.,
RA Winkler P., Lander B.S., Weissenbach J., Roest Crolius H.;
RT "genome duplication in the teleost fish Tetraodon nigroviridis reveals
RT the early vertebrate proto-karyotype.";
RL Nature 431:946-957(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RG Genoscope; Whitehead Institute Centre for Genome Research;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -! CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC -! FUNCTION: Plays a key role in the control of the eukaryotic cell
CC cycle. It is required in higher cells for entry into S-phase and
CC mitosis. Component of the kinase complex that phosphorylates the
CC repetitive C-terminus of RNA polymerase II. Catalytic component of
CC MPF (By similarity).

```

```

CC -! SUBUNIT: Forms a stable but non-covalent complex with cyclin B in
CC mature oocytes (By similarity).
CC EMBL: CAAE01014700; CAG02796.1; -; Genomic DNA.
DR InterPro: IPR010472; DUF_FH3.
DR InterPro: IPR010473; DUF_GTPase_bd.
DR InterPro: IPR003104; FH2_actin_bd.
DR InterPro: IPR001265; Formin.
DR Pfam: PF06367; DUF_FH3; 1.
DR Pfam: PF06371; DUF_GBD; 1.
DR Pfam: PF02181; FH2; 1.
DR PRINTS: PRO0828; FORMIN.
DR SMART: SM00498; FH2; 1.
FT NON TER 1
SQ SEQUENCE 1204 AA; 134434 MW; 7468A1301977D100 CRC64;

Query Match 13.0%; Score 101; DB 2; Length 1204;
Best Local Similarity 33.3%; Pred. No. 16;
Matches 38; Conservative 2; Mismatches 32; Indels 42; Gaps 5;

QY 4 NDDPAPGSDASRSR--PAPALG-----PPPPASGWFDSGLVSPRPICAASSAGLP 52
DB 584 NLSPAPASDVCDASAPPPPALGVPVPPPPPPPPPPPPALGANGA---PPPPPPPPSAGLP 640

QY 53 PPVPTPTLVNDVTCGSGWVSCIGPLISPMRVWVAAGNPTGVLPGEIGI 106
DB 641 PPPPP-----PL-----PQAGPPPPPPPPPPPPGAGFP 666

RESULT 10
Q6ZVX4_HUMAN
ID Q6ZVX4_HUMAN PRELIMINARY; PRT; 132 AA.
AC Q6ZVX4;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein FLJ41974.
DS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Oshima A., Takahashi-Fuji A., Tanase T., Imose N., Takeuchi K.,
RA Arita M., Muraeshino K., Yuuki H., Hara H., Sugiyama T., Irie R.,
RA Otsuki T., Sato H., Ota T., Wakamatsu A., Ishii S., Yamamoto J.,
RA Isono Y., Kawai-Hio Y., Saito K., Nishikawa T., Kimura K.,
RA Yamashita H., Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K.,
RA Wagaatsuma M., Murakawa K., Kanehori K., Sugiyama A., Kawakami B.,
RA Suzuki Y., Sugano S., Nagahara K., Masuhio Y., Nagai K., Isegai T.;
RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL: AK123968; BAC85736.1; -; mRNA.
SQ SEQUENCE 132 AA; 13442 MW; 2142A7AD97B633D CRC64;

Query Match 12.8%; Score 99; DB 2; Length 132;
Best Local Similarity 33.7%; Pred. No. 2.8;
Matches 35; Conservative 6; Mismatches 41; Indels 22; Gaps 7;

QY 6 PPAPGSDASRSRPPAPALGPPPPAS-GWFDGGLVSPRPICAASSAGLPVPPPTLVNDV 64
DB 19 PTEG-----AAVTSPSSPPQASLAWRSSG--SRPCCPLTSLTSP-----WRNQL 64

QY 65 -TCCSGWVSCIGPLISPSW-----PRVWVAAGNPTGVLP 102
DB 65 HTAPSGGCGCTVPAGSIAWAGSPSRMTFQGLRATSR-LPG 107

RESULT 11
Q4S68_TETNG
ID Q4S68_TETNG PRELIMINARY; PRT; 494 AA.
AC Q4S68;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)

```



```

DT 13-SRP-2005 (TrEMBLrel. 31, last annotation update)
DE Chromosome undetermined SCAFI4724, whole genome shotgun sequence.
DN (Fragment)
GN ORFNames=GSTENG00023156001;
OS Tetraodon nigroviridis (green puffer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Tetraodon.
OX NCBI_TaxID=99883;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Jallion O., Aury J.M., Brunet P., Petit J.L., Stange-Thomann N.,
RA Mauceli E., Bouneau L., Fischer C., Ozouf-Costet C., Bernot A.,
RA Nicard S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
RA Daeilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
RA Anthouard V., Jubin C., Castelli V., Katinka M., Vacherie B.,
RA Bismont C., Skallil Z., Catcolico L., Pouliat J., De Bernardis V.,
RA Craud C., Duprat S., Broctier P., Coutanceau J.P., Guzy J.,
RA Parra G., Lardier G., Chapple C., McKernan K.J., McEwan P., Bosak S.,
RA Kellis M., Voltl J.N., Gilgo R., Zody M.C., Westro J.,
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA Landet V., Schachter V., Queller F., Sautin W., Scarpelli C.,
RA Winkler P., Lander E.S., Weissbach J., Roest Crolians H.;
RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
RT the early vertebrate proto-karyotype."
RL Nature 431:946-957(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RG Genoscope; Whitehead Institute Centre for Genome Research;
RL Submitted (FEB-2004) to the EMBL/Genbank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/Genbank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; CAAB01044724; CAG03664.1; -; Genomic_DNA.
FT NON TER 1
SQ SEQUENCE 494 AA; 55841 MW; 2F0AE1BA8BFCD3B8 CRC64;

Query Match 12.8%; Score 99; DB 2; Length 494;
Best Local Similarity 29.7%; Pred. No. 9.8;
Matches 41; Conservative 10; Mismatches 31; Indels 56; Gaps 7;

QY 4 NDPAPRGDSARSRAPALGPPPPASGMPDGLVSRP1-----CASSSAGL 51
DB 153 SERPADGDSAFSRP-PALHQPAPA-----LPPKPLNRLPNH/TKLP--MSMKL 200
QY 52 PEPVPEPTLNNDVYCCSGMVCCLGPLSPS-----WPRVWVAAGNPTVGLPGE 103
DB 201 SPRLPPKTL-----MISVPVPSALTFQKCPAPSHUPAGSHPOYGALPAP 247
QY 104 GIPKGFVLMILAPGSR1 121
DB 248 -----LHPSPRI 254

RESULT 12
06ZCX3 ORYSA PRELIMINARY; PRT; 893 AA.
AC 06ZCX3;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Putative diaphanous 1.
GN Name=P0026A08.31-1; Synonyms=OJ1003.B05.5-1;
OS Oryza sativa (Japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzoae; Oryza.
NCBI_TaxID=39947;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Sasaki T., Matsumoto T., Yamamoto K.;
RL Submitted (DEC-2001) to the EMBL/Genbank/DBJ databases.

```

[illegible]

Qy 122 DAIGSFSKSVLTAV 136
Db 121 DVGAGPDYGPIMPAV 135

RESULT 14

Q8VK97_MYCTU
ID Q8VK97_MYCTU PRELIMINARY; PRT; 437 AA.
AC Q8VK97;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, last annotation update)
DE Hypothetical protein.
GN OrderedLocustNames=MT1033;
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium;
OC Mycobacterium tuberculosis complex.
OX NCBI_TaxID=1773;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CDC 1551 / Oehkosh;
RX MEDLINE=22206494; PubMed=12218036;
RX DOI=10.1128/JB.184.19.5479-5490.2002;
RA Fieischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J.D., DeBoy R.T., Dodson R.J., Gwinn M.L., Haft D.H.,
RA Hickey B.K., Kolonay J.F., Nelson W.C., Umayan L.A., Ermolaeva M.D.,
RA Salzberg S.L., Delcher A., Utterback T.R., Weidman J.F., Khouri H.M.,
RA Gill J., Mikula A., Bishai W., Jacobs W.R. Jr., Venter J.C.,
RA Fraser C.M.;
RT "Whole-genome comparison of Mycobacterium tuberculosis clinical and
laboratory strains."
RL J. Bacteriol. 184:5479-5490 (2002).
DR EMBL; AB000516; AKA45283.1; -; Genomic_DNA.
DR TIGR; MT1033; -.
KW Hypothetical protein.
SQ SEQUENCE 437 AA; 40687 MW; 35934558B8706B89 CRC64;

Query Match 12.7%; Score 98.5; DB 2; Length 437;

Best Local Similarity 29.3%; Pred. No. 9.5;
Matches 39; Conservative 13; Mismatches 62; Indels 19; Gaps 5;

Qy 5 DPPAGSDASRPAALGPDPPASGWFDSGLVPSRPTCAASSSAGLPPVPPTMLNDV 64
Db 77 DTATPSPGPAAEHVPVPA--PPRPVADRTIAPGLVPRPGVPAALAAVRCPPAVPGLPN--- 131
Qy 65 TCCSGWVSCCTGPIISPSRWVVAAGNMPTGVELPGEIGPKIGFVVLMAPGSRIDAI 124
Db 132 -----IPGAAGPALPPPP-PALPNMLAASVP-GVGIPGIGIGI-----PGIGIPGI 177
Qy 125 GSFSFSKSVLTAVS 137
Db 178 GIPGVPPPTIGVN 190

RESULT 15

059GF6_HUMAN

ID 059GF6_HUMAN PRELIMINARY; PRT; 1332 AA.

AC 059GF6;
DT 10-MAY-2005 (TREMBlrel. 30, Created)
DT 10-MAY-2005 (TREMBlrel. 30, last sequence update)
DT 10-MAY-2005 (TREMBlrel. 30, last annotation update)
DE Formin 2 variant (Fragment).
GN Name=Formin 2 variant;
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Brain;

RA Totoki Y., Toyoda A., Takeda T., Sakaki Y., Tanaka A., Yokoyama S.,
RA Ohara O., Nagase T., Kikuno F.R.;
RT "None Title."
RL Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB209153; BAD92390.1; -; mRNA.
FT NON TER 1
SQ SEQUENCE 1332 AA; 140520 MW; DE3DA9FBB1445DB CRC64;

Query Match 12.7%; Score 98.5; DB 2; Length 1332;
Best Local Similarity 26.0%; Pred. No. 28;
Matches 38; Conservative 10; Mismatches 47; Indels 51; Gaps 7;

Qy 6 PPPAGSDASRPAALG--PDPPASGWFDSGLVPSRPTCAASSSAGLPPV-----VP 56
Db 743 PPLPGAGIPPPPLPGVGIPPPPLPG--AGIPPPPL----PGAGIPPPPLPGVGIP 795
Qy 57 PTMLNDVTCGSGVSCCTGPIISPSRWVW-----AAGNMPTGVELPGEIGPKIGFV 111
Db 796 P-----PPPLGVGIPPPPLPGAGIPPPPLPGMGIIP----- 829
Qy 112 VMLAPGSRIDAISSFSKSVLTAVS 137
Db 830 -----APAPPLPPPGTGI PPPPLPV 851

Search completed: April 14, 2006, 17:32:28
Job time : 54.2864 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using SW model

Run on: April 14, 2006, 17:32:59 ; Search time 12.853 Seconds
(without alignments)
894.107 Million cell updates/sec

Title: US-10-620-246-90

Perfect score: 774
Sequence: 1 MRVNDPPAPGSDARSRRPAP.....RIDALGSSFSKSVLTAVASAW 139

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: Issued_Patents_AA.*
2: /cgn2_6/ptodata/1/1aa/5.COMB.pep.*
3: /cgn2_6/ptodata/1/1aa/6.COMB.pep.*
4: /cgn2_6/ptodata/1/1aa/H.COMB.pep.*
5: /cgn2_6/ptodata/1/1aa/PCITUS.COMB.pep.*
6: /cgn2_6/ptodata/1/1aa/BACKL1EET.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	774	100.0	139	2	US-09-050-739-90 Sequence 90, Appl
2	103.5	13.4	707	2	US-09-919-039-778 Sequence 278, App
3	103.5	13.4	707	2	US-09-538-092-993 Sequence 993, App
4	103.5	13.4	735	2	US-09-949-016-10120 Sequence 10120, A
5	101	13.0	273	2	US-09-252-991A-28324 Sequence 28324, A
6	99.5	12.9	198	2	US-09-252-991A-28328 Sequence 28328, A
7	96	12.4	663	2	US-09-252-991A-30843 Sequence 30843, A
8	95.5	12.3	175	2	US-09-252-991A-21648 Sequence 21648, A
9	95.5	12.3	794	2	US-09-252-991A-28569 Sequence 28569, A
10	95	12.3	313	2	US-09-270-767-43688 Sequence 43688, A
11	93.5	12.1	379	2	US-09-949-016-10372 Sequence 10372, A
12	93	12.0	442	2	US-09-252-991A-23285 Sequence 23285, A
13	92.5	12.0	187	2	US-09-199-637A-287 Sequence 287, App
14	92.5	12.0	187	2	US-09-252-991A-21454 Sequence 21454, A
15	91.5	11.8	414	2	US-09-252-991A-25096 Sequence 25096, A
16	91.5	11.8	502	2	US-09-839-577A-14 Sequence 14, Appl
17	91	11.8	139	2	US-09-252-991A-27651 Sequence 27651, A
18	90.5	11.7	380	1	US-09-026-587-4 Sequence 4, Appl1
19	90.5	11.7	380	1	US-09-227-420-4 Sequence 4, Appl1
20	90.5	11.7	380	2	US-09-387-811-4 Sequence 4, Appl1
21	90.5	11.7	380	2	US-09-823-240A-10 Sequence 10, Appl
22	90.5	11.7	381	2	US-09-949-016-7291 Sequence 7291, App
23	90.5	11.7	581	2	US-09-949-016-9978 Sequence 9978, App
24	90.5	11.7	635	2	US-09-417-197-125 Sequence 125, App
25	90	11.6	148	2	US-09-252-991A-23769 Sequence 23769, A
26	90	11.6	158	2	US-09-252-991A-23664 Sequence 23664, A
27	89.5	11.6	65	2	US-09-252-991A-28226 Sequence 28226, A

28	89	11.5	137	2	US-09-489-039A-12891 Sequence 12891, A
29	89	11.5	489	2	US-09-518-914-4 Sequence 4, Appl1
30	89	11.5	498	1	US-08-660-963-12 Sequence 12, Appl
31	89	11.5	503	2	US-09-599-287A-2 Sequence 2, Appl1
32	89	11.5	503	2	US-10-078-547-2 Sequence 2, Appl1
33	89	11.5	506	2	US-09-949-016-11282 Sequence 11282, A
34	89	11.5	507	2	US-09-599-287A-24 Sequence 24, Appl
35	89	11.5	507	2	US-10-078-547-24 Sequence 24, Appl
36	88.5	11.4	191	2	US-09-252-991A-19331 Sequence 19331, A
37	88.5	11.4	1248	2	US-09-080-897-2 Sequence 2, Appl1
38	88.5	11.4	1248	2	US-09-323-735-2 Sequence 2, Appl1
39	88.5	11.4	1315	2	US-08-899-595-3 Sequence 3, Appl1
40	88	11.4	478	2	US-09-518-914-2 Sequence 2, Appl1
41	87.5	11.3	162	2	US-09-252-991A-32485 Sequence 32485, A
42	87.5	11.3	211	2	US-09-252-991A-24927 Sequence 24927, A
43	87.5	11.3	225	2	US-09-252-991A-30156 Sequence 30156, A
44	87.5	11.3	308	2	US-09-252-991A-18097 Sequence 18097, A
45	87	11.2	415	2	US-09-252-991A-20613 Sequence 20613, A

ALIGNMENTS

RESULT 1
US-09-050-739-90
Sequence 90, Application US/09050739
Patent No. 6641814
GENERAL INFORMATION:
APPLICANT: ANDERSEN, Peter
APPLICANT: NIELSEN, Rikke
APPLICANT: OETTINGER, Thomas
APPLICANT: RASMUSSEN, Peter Birk
APPLICANT: ROSENKRANDS, Ida
APPLICANT: WEIDINGH, Karin
APPLICANT: FLOBO, Walter
TITLE OF INVENTION: NUCLEIC ACIDS FRAGMENTS AND POLYPEPTIDE FRAGMENTS
FILE REFERENCE: 670001-2002.1
CURRENT APPLICATION NUMBER: US/09/050, 739
CURRENT FILING DATE: 1998-03-30
EARLIER APPLICATION NUMBER: 0376/97
EARLIER FILING DATE: 1997-04-02
EARLIER APPLICATION NUMBER: 1277/97
EARLIER FILING DATE: 1997-11-10
EARLIER APPLICATION NUMBER: 60/044, 624
EARLIER FILING DATE: 1997-04-18
EARLIER APPLICATION NUMBER: 60/070, 488
EARLIER FILING DATE: 1998-01-05
NUMBER OF SEQ ID NOS: 173
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 90
LENGTH: 139
TYPE: PRT
ORGANISM: Mycobacterium tuberculosis
US-09-050-739-90

Query Match 100.0%; Score 774; DB 2; Length 139;
Best Local Similarity 100.0%; Pred. No. 2.3e-66;
Matches 139; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MRVNDPPAPGSDARSRRPAPALGPDPPASGMFDSGLVPSRPTCAASSAGLPPVPPTWL	60
DB	1	MRVNDPPAPGSDARSRRPAPALGPDPPASGMFDSGLVPSRPTCAASSAGLPPVPPTWL	60
QY	61	NNDTCCSGVNSCCIGLISPSWPRVVAAGMPTVEVLPGSGIPKIGFVTLAQR	120
DB	61	NNDTCCSGVNSCCIGLISPSWPRVVAAGMPTVEVLPGSGIPKIGFVTLAQR	120
QY	121	IDALGSSFSKSVLTAVASAW	139
DB	121	IDALGSSFSKSVLTAVASAW	139

US-09-252-991A-28324

Query Match 13.0%; Score 101; DB 2; Length 273;
 Best Local Similarity 29.6%; Pred. No. 0.064;
 Matches 40; Conservative 7; Mismatches 48; Indels 40; Gaps 5;

QY 6 PPAGSDSARSPPALGDPDPAGWFDGSLVPSRPICAASSSAGLPPVPTWLANDVT 45
 DB 10 PPRTTAAAWPMPADSA.PATGSGPAAAGRHSLPPPARSIRRVRRACAGRVASAPG 69
 QY 46 SSSAGLPPVPTWLANDVTCGSGWVSCCTGPLISPSRPVVAAGWPTGVELPGEGI 105
 DB 70 SSRARAPGAPPR-----GCGRAMPAGWMTSAAN---GSPRSCRC 109
 QY 106 PKIGFVVLMLAPGSR 120
 DB 110 ATAGRVPAASASGR 124

RESULT 6

US-09-252-991A-28328
 ; Sequence 28328, Application US/09252991A
 ; Patent No. 6551795

; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252.991A
 ; CURRENT FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 28328
 ; LENGTH: 198
 ; TYPE: PRT
 ; ORGANISM: Pseudomonas aeruginosa
 US-09-252-991A-28328

Query Match 12.9%; Score 99.5; DB 2; Length 198;
 Best Local Similarity 27.7%; Pred. No. 0.062;
 Matches 31; Conservative 12; Mismatches 38; Indels 31; Gaps 5;

QY 6 PPAGSDSARSPPALGDPDPAGWFDGSLVPSRPICAASSSAGLPPVPTWLANDVT 65
 DB 36 PPATWTASPSRPTSPAPPAPAT-----PRRIPPSACSPGAPAP-----RRAS 80
 QY 66 CCGWVSC---CIGPLISPS-WPRVW-----VAAGWPTGVELP 101
 DB 81 AATTWKAACGRCGSAATSHWPNWARSAPSCWMSATSIKACWMPSSSSVP 132

RESULT 7

US-09-252-991A-30843
 ; Sequence 30843, Application US/09252991A
 ; Patent No. 6551795

; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252.991A
 ; CURRENT FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 30843
 ; LENGTH: 663
 ; TYPE: PRT

; ORGANISM: Pseudomonas aeruginosa
 US-09-252-991A-30843

Query Match 12.4%; Score 96; DB 2; Length 663;
 Best Local Similarity 27.8%; Pred. No. 0.53;
 Matches 37; Conservative 9; Mismatches 45; Indels 42; Gaps 8;

QY 6 PPAGSDSARSPPALGDPDPAGWFDGSLVPSRPICAASSSAGLPPVPTWLANDVT 62
 DB 30 PPADRSATARVAPACRVPTRCANAW-----PGRPMSTSRPGRWSPAVPAGACWKTR 83
 QY 63 DV---TCCSGWV-----CCT---GPLISPSRP-----RVVAAGWPTGVELPGEGI 104
 DB 84 NVAGRSATDAMSSAPAPANCCLPPAGP--SPAPADACRWRRAACHWPAS----- 134
 QY 105 IPKIGFVVLMLAP 117
 DB 135 -----AMWMLAP 141

RESULT 8

US-09-252-991A-21648
 ; Sequence 21648, Application US/09252991A
 ; Patent No. 6551795

; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252.991A
 ; CURRENT FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 21648
 ; LENGTH: 175
 ; TYPE: PRT
 ; ORGANISM: Pseudomonas aeruginosa
 US-09-252-991A-21648

Query Match 12.3%; Score 95.5; DB 2; Length 175;
 Best Local Similarity 22.0%; Pred. No. 0.13;
 Matches 28; Conservative 13; Mismatches 37; Indels 49; Gaps 5;

QY 7 PPAGSDSARSPPALGDPDPAGW-----FDSGLVPSRP 41
 DB 4 PSTVPPKASRHSQSPASPASAWSCRDVPSMSTCARGTCTSPYRAGCWYASPPCRS 63
 QY 42 ICAASSAGL-----PPVPTWLANDVTCGSGWVSC---CIGPLISPSWPRVWV 88
 DB 64 SCSTSATTPMPSSKACRBPAPP-----CSAGWTCPRASSGALCTGMA---A 112

QY 89 AAGGNWP 95
 DB 113 SASSPWP 119

RESULT 9

US-09-252-991A-28569
 ; Sequence 28569, Application US/09252991A
 ; Patent No. 6551795

; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252.991A
 ; CURRENT FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190

```
/ PRIOR FILING DATE: 1998-07-27
/ NUMBER OF SEQ ID NOS: 33142
/ SEQ ID NO 28569
/ LENGTH: 794
/ TYPE: PRT
/ ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-28569
```

```
Query Match
Best Local Similarity 12.3%; Score 95.5; DB 2; Length 794;
Pred. No. 0.73;
Matches 44; Conservative 18; Mismatches 59; Indels 43; Gaps 10;
```

```
QY 3 VNDPPAPG-----SDSARSRP--APALGDDPPASGWFDSG-----LVSPRPICA-- 44
DB 81 VNRPRPSGCGNCRSTPRASTTSGSRPICAPAI-----ASRWAGSARCGSTPTAPATPAAR 135
QY 45 ASSSAGLPPVPTWLNNDVTCGSGW-VSCCIGPLISPRVVAAGNW-----P 95
DB 136 AASAASTPWPAPMPWATNRGRSASPWPTSVSMASRSRTTPSATGAPAPRPWSPTRWATP 195
QY 96 TGVE---LPGEGTPKIGFVVLMLAPGRIDAGSSPSKSVLTAV 136
DB 196 TSTDRAASVPG-----NYANTWTS--AKSACLTGSAKRCMPAAV 231
```

RESULT 10

```
US-09-270-767-43688
/ Sequence 43688, Application US/09270767
/ Patent No. 6703491
/ GENERAL INFORMATION:
/ APPLICANT: Homburger et al.
/ TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
/ FILE REFERENCE: File Reference: 7326-094
/ CURRENT APPLICATION NUMBER: US/09/270,767
/ NUMBER OF SEQ ID NOS: 62517
/ SOFTWARE: PatentIn Ver. 2.0
/ LENGTH: 313
/ TYPE: PRT
/ ORGANISM: Drosophila melanogaster
US-09-270-767-43688
```

```
Query Match
Best Local Similarity 12.3%; Score 95; DB 2; Length 313;
Pred. No. 0.28;
Matches 43; Conservative 16; Mismatches 61; Indels 46; Gaps 8;
```

```
QY 7 PARGSDSARSRPA-----PALGDDPPASGWFDSGLVSPRPICT--AASSAGLPP----- 54
DB 23 PARGAPAGSNVVPAGSSPSGAVLGNPPA-----GSVPSIPVLPGAGSWPMHPRPIPI 77
QY 55 VPTTW-----LNDVTCGSGWVSCCIGP--LISPSRVVVAAGNW----- 94
DB 78 VPTWRRLPLRGQSGAGPGSGGLINGINVPQRPHPFMPN--WQSVNWSMRPTKVPVSTEAR 136
QY 95 -----PTGVELPGEIGPKIGFVVLMLAPGRIDAGSSPSKSV 132
DB 137 VQPNPQGPQNPQPGNSNPETASVQAAFPVPANVPANENASV 182
```

RESULT 11

```
US-09-949-016-10372
/ Sequence 10372, Application US/09949016
/ Patent No. 6812339
/ GENERAL INFORMATION:
/ APPLICANT: VENTER, J. Craig et al.
/ TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
/ TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
/ FILE REFERENCE: C1001307
/ CURRENT APPLICATION NUMBER: US/09/949,016
/ PRIOR FILING DATE: 2000-04-14
/ PRIOR APPLICATION NUMBER: 60/241,755
/ PRIOR FILING DATE: 2000-10-20
```

```
/ PRIOR APPLICATION NUMBER: 60/237,768
/ PRIOR FILING DATE: 2000-10-03
/ PRIOR APPLICATION NUMBER: 60/231,498
/ PRIOR FILING DATE: 2000-09-08
/ NUMBER OF SEQ ID NOS: 207012
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 10372
/ LENGTH: 379
/ TYPE: PRT
/ ORGANISM: Human
US-09-949-016-10372
```

```
Query Match
Best Local Similarity 12.1%; Score 93.5; DB 2; Length 379;
Pred. No. 0.49;
Matches 36; Conservative 4; Mismatches 39; Indels 31; Gaps 6;
```

```
QY 6 PPARGSDSARSRPAPALGDDPPASGWFDSGLVSPRPICAASSAGL--PPVPTWLNNDV 64
DB 269 PPVGAEMAMARVYQPPAPPRP-----PQSASSLSASLRPPGAPATFLRPSP 315
QY 65 TCCSG---WVSCC--IGPLISPSWPRVVAAGNWPT---GVBLPGEGIPK 107
DB 316 IPCSSPQPNQSLGGLP-----PCAGCPWPTAGPGRBRPGGTSPR 355
```

RESULT 12

```
US-09-252-991A-23285
/ Sequence 23285, Application US/09252991A
/ Patent No. 6551795
/ GENERAL INFORMATION:
/ APPLICANT: Marc J. Rubenfield et al.
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
/ TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
/ FILE REFERENCE: 107196.136
/ CURRENT APPLICATION NUMBER: US/09/252,991A
/ PRIOR FILING DATE: 1999-02-18
/ PRIOR APPLICATION NUMBER: US 60/074,788
/ PRIOR FILING DATE: 1998-02-18
/ PRIOR APPLICATION NUMBER: US 60/094,190
/ PRIOR FILING DATE: 1998-07-27
/ NUMBER OF SEQ ID NOS: 33142
/ SEQ ID NO 23285
/ LENGTH: 442
/ TYPE: PRT
/ ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-23285
```

```
Query Match
Best Local Similarity 12.0%; Score 93; DB 2; Length 442;
Pred. No. 0.65;
Matches 47; Conservative 9; Mismatches 61; Indels 72; Gaps 9;
```

```
QY 6 PP-----ARGSDSARSRPAPALGDDPPASGWFDSGLVSPRPICAASSAGLPPVPTW 59
DB 32 PPLAAGCAAPG--RAARYPPTATGRSAPAGWSPPCRSRAVLQAIRSPAPRP----- 84
QY 60 LNDVTCGS-----GWVSCCIGPLISPS 82
DB 85 --GTVSSCSRSRDLPTAGTRESRAHSFSGAGRHAGARRSRPAEFGW-----RPGSRPG 137
QY 83 WP--RYVVA-----AGNMPTGVLEP-----GEGIPKIGFVVLMLAPGRIDAGSSFS 129
DB 138 WPARRGWSRAVRPSLARODPPACPRPRPCRTGAABRRITGFARAWFT--SPAPAGASLA 195
QY 130 KSVLTAVSA 138
DB 196 IAVCIRPASA 204
```

RESULT 13

```
US-09-199-637A-287
/ Sequence 287, Application US/09199637A
/ Patent No. 6355411
/ GENERAL INFORMATION:
```

```
; APPLICANT: Aueubel, Frederick
; APPLICANT: Goodman, Howard M.
; APPLICANT: Rahme, Laurence G.
; APPLICANT: Mahajan-Miklos, Shalina
; APPLICANT: Tan, Man-Wah
; APPLICANT: Cao, Hui
; APPLICANT: Drenkard, Eliana
; APPLICANT: Teonggalib, John
; TITLE OF INVENTION: VIRULENCE-ASSOCIATED NUCLEIC ACID
; TITLE OF INVENTION: SEQUENCES AND USES THEREOF
; FILE REFERENCE: 00786/361002
; CURRENT APPLICATION NUMBER: US/09/199,637A
; PRIOR FILING DATE: 1998-11-25
; PRIOR APPLICATION NUMBER: 60/066,517
; NUMBER OF SEQ ID NOS: 437
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 287
; LENGTH: 187
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-199-637A-287
```

```
Query Match 12.0%; Score 92.5; DB 2; Length 187;
Best Local Similarity 32.5%; Pred. No. 0.27;
Matches 37; Conservative 7; Mismatches 49; Indels 21; Gaps 8;
```

```
QY 6 PPARGSDSARSRPAPALGPD---PPASG-----WFDGLVPSR-PICAASSS---AG 50
DB 61 PPRGRSLADAPAPCAASARPRASSTPGYRSTWKLRRFPRRSPCCATGSRSCNAG 120

QY 51 LPP-VP-PTWLNNDVTCGSGWVSCCIGPLISPSWPRVWAAGNMPTGVLEP 101
DB 121 RPSAGCVPGMGWGTGCAPCSRGRPCA-APPSPAMPA--RSSAGSAPSPAPRP 171
```

```
RESULT 14
US-09-252-991A-21454
; Sequence 21454, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 21454
; LENGTH: 187
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-21454
```

```
Query Match 12.0%; Score 92.5; DB 2; Length 187;
Best Local Similarity 32.5%; Pred. No. 0.27;
Matches 37; Conservative 7; Mismatches 49; Indels 21; Gaps 8;
```

```
QY 6 PPARGSDSARSRPAPALGPD---PPASG-----WFDGLVPSR-PICAASSS---AG 50
DB 61 PPRGRSLADAPAPCAASARPRASSTPGYRSTWKLRRFPRRSPCCATGSRSCNAG 120

QY 51 LPP-VP-PTWLNNDVTCGSGWVSCCIGPLISPSWPRVWAAGNMPTGVLEP 101
DB 121 RPSAGCVPGMGWGTGCAPCSRGRPCA-APPSPAMPA--RSSAGSAPSPAPRP 171
```

```
RESULT 15
US-09-252-991A-25096
```

```
; Sequence 25096, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 25096
; LENGTH: 414
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-25096
```

```
Query Match 11.8%; Score 91.5; DB 2; Length 414;
Best Local Similarity 35.6%; Pred. No. 0.83;
Matches 26; Conservative 6; Mismatches 26; Indels 15; Gaps 3;
```

```
QY 7 PAPGSDSARSRPAPALGPDPPASGM-----FDGLVPSRPTICAASSAGLPPVPTWL 60
DB 20 PAACSNRARCRRPMPG---PAAGWSTTAPWPSRRPSGPICQASSSTRRSPTATAS-- 74

QY 61 NNDVTCGSGWVSC 73
DB 75 ---AVCSRWTTC 83
```

```
Search completed: April 14, 2006, 17:37:59
Job time : 13.853 secs
```

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using SW model

Run on: April 14, 2006, 18:27:13 ; Search time 48.7785 Seconds
(without alignments)
1190.652 Million cell updates/sec

Title: US-10-620-246-90

Perfect score: 774

Sequence: 1 MRVNDPPAPGSDARSRRPAP.....RIDAGSFSKSVLTAVSAW 139

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_AAMain:*

- 1: /cgn2_6/ptodaca/1/pubpaa/US07_PUBCOMB.pep:*
- 2: /cgn2_6/ptodaca/1/pubpaa/US08_PUBCOMB.pep:*
- 3: /cgn2_6/ptodaca/1/pubpaa/US09_PUBCOMB.pep:*
- 4: /cgn2_6/ptodaca/1/pubpaa/US10_PUBCOMB.pep:*
- 5: /cgn2_6/ptodaca/1/pubpaa/US10B_PUBCOMB.pep:*
- 6: /cgn2_6/ptodaca/1/pubpaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	774	100.0	139	US-09-791-171-90	Sequence 90, Appl
2	774	100.0	139	US-09-804-980-90	Sequence 90, Appl
3	774	100.0	139	US-10-620-246-90	Sequence 90, Appl
4	112	14.5	377	US-10-437-963-133141	Sequence 133141,
5	110	14.2	123	US-10-767-701-53489	Sequence 53489, A
6	103.5	13.4	707	US-09-919-039-278	Sequence 278, App
7	102	13.2	135	US-09-864-761-33360	Sequence 33360, A
8	101.5	13.1	285	US-10-437-963-185237	Sequence 185237,
9	99.5	12.9	179	US-10-437-963-169481	Sequence 169481,
10	99.5	12.9	199	US-10-029-386-33093	Sequence 33093, A
11	99.5	12.9	212	US-09-925-302-741	Sequence 741, App
12	99.5	12.9	212	US-09-925-302-741	Sequence 741, App
13	99.5	12.9	297	US-10-437-963-163320	Sequence 163320,
14	99	12.8	327	US-10-437-963-114935	Sequence 114935,
15	98	12.7	155	US-10-767-701-40674	Sequence 40674, A
16	98	12.7	457	US-10-437-963-141396	Sequence 141396,
17	97.5	12.6	252	US-10-437-963-151853	Sequence 151853,
18	97.5	12.6	19662	US-10-084-846A-6	Sequence 6, Appl1
19	97	12.5	198	US-10-029-386-33094	Sequence 33094, A
20	96.5	12.5	354	US-10-437-963-163319	Sequence 163319,
21	96.5	12.5	196	US-10-425-115-292580	Sequence 292580,
22	96.5	12.5	304	US-10-437-963-192758	Sequence 192758,
23	96.5	12.5	1049	US-10-317-835-16	Sequence 16, Appl
24	96.5	12.5	1049	US-10-168-582-7	Sequence 7, Appl1
25	96	12.4	170	US-10-767-701-36434	Sequence 36434, A
26	96	12.4	387	US-10-450-763-54234	Sequence 54234, A
27	96	12.4	1507	US-10-437-963-143963	Sequence 143963,

28	96	12.4	1812	4	US-10-437-963-151404	Sequence 151404,
29	95	12.3	295	4	US-10-425-115-283342	Sequence 283342,
30	95	12.3	1612	6	US-11-097-143-5022	Sequence 5022, Ap
31	95	12.3	19725	4	US-10-084-846A-4	Sequence 4, Appl1
32	94.5	12.2	254	4	US-10-437-963-173295	Sequence 173295,
33	94	12.1	149	4	US-10-437-963-204838	Sequence 204838,
34	94	12.1	157	4	US-10-437-963-156311	Sequence 156311,
35	94	12.1	198	4	US-10-767-701-58165	Sequence 58165, A
36	94	12.1	460	5	US-10-739-930-8672	Sequence 8672, Ap
37	94	12.1	475	4	US-10-437-963-187947	Sequence 187947,
38	94	12.1	632	4	US-10-282-122A-51472	Sequence 51472, A
39	94	12.1	19695	4	US-10-084-846A-3	Sequence 3, Appl1
40	93.5	12.1	217	4	US-10-767-701-37061	Sequence 37061, A
41	93.5	12.1	240	4	US-10-437-963-144769	Sequence 144769,
42	93.5	12.1	365	3	US-09-876-187-4	Sequence 4, Appl1
43	93.5	12.1	365	3	US-09-749-728B-15	Sequence 15, Appl
44	93.5	12.1	365	5	US-10-732-923-16081	Sequence 16081, A
45	93.5	12.1	914	5	US-10-450-763-57757	Sequence 57757, A

ALIGNMENTS

RESULT 1						
US-09-791-171-90						
Sequence 90, Application US/09791171						
Patent No. US20020094336A1						
GENERAL INFORMATION:						
APPLICANT: ANDERSEN, Peter						
APPLICANT: NIELSEN, Rikke						
APPLICANT: OETTINGER, Thomas						
APPLICANT: RASMUSSEN, Peter Blak						
APPLICANT: ROSENKRANDS, Ida						
APPLICANT: WELDMING, Karin						
APPLICANT: FLORIO, Walter						
TITLE OR INVENTION: NUCLEIC ACIDS FRAGMENTS AND POLYPEPTIDE FRAGMENTS						
FILE REFERENCE: 670001-2002.1						
CURRENT APPLICATION NUMBER: US/09/791,171						
CURRENT FILING DATE: 2001-02-20						
PRIOR APPLICATION NUMBER: 09/050,739						
PRIOR FILING DATE: 1998-03-30						
PRIOR APPLICATION NUMBER: 0376/97						
PRIOR FILING DATE: 1997-04-02						
PRIOR APPLICATION NUMBER: 1277/97						
PRIOR FILING DATE: 1997-11-10						
PRIOR APPLICATION NUMBER: 60/044,624						
PRIOR FILING DATE: 1997-04-18						
PRIOR APPLICATION NUMBER: 60/070,488						
PRIOR FILING DATE: 1998-01-05						
NUMBER OF SEQ ID NOS: 173						
SOFTWARE: PatentIn Ver. 2.0						
SEQ ID NO 90						
LENGTH: 139						
TYPE: PRT						
ORGANISM: Mycobacterium tuberculosis						
US-09-791-171-90						
Query Match						
Best local similarity 100.0%; Score 774; DB 3; Length 139;						
Matches 139; Conservative 0; Mismatches 0; Indels 0; Gaps 0;						
QY	1	MRVNDPPAPGSDARSRRPAPALGDDPPASGMFSGVLSRPRICASSAGLPVPPTWL	60			
DB	1	MRVNDPPAPGSDARSRRPAPALGDDPPASGMFSGVLSRPRICASSAGLPVPPTWL	60			
QY	61	NNDVTCGSGVSCICIGLISPSWPRVWVAAGNMPTGVELPGGCIPIRIGFVVIWLABGR	120			
DB	61	NNDVTCGSGVSCICIGLISPSWPRVWVAAGNMPTGVELPGGCIPIRIGFVVIWLABGR	120			
QY	121	IDAIGSFSKSVLTAVSAW 139				
DB	121	IDAIGSFSKSVLTAVSAW 139				

PRIOR FILING DATE: 1999-01-21

RESULT 5
US-10-767-701-53499

RESULT 5

Sequence 53499, Application US/10767701
Publication No. US20040172684A1
GENERAL INFORMATION:
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53535)B
CURRENT FILING DATE: US/10/767,701
NUMBER OF SEQ ID NOS: 63128
SEQ ID NO 53499
LENGTH: 123
TYPE: PRT
ORGANISM: Sorghum bicolor
FEATURE:
NAME/KEY: unsure
LOCATION: (1)..(123)
OTHER INFORMATION: unsure at all Xaa locations
FEATURE:
OTHER INFORMATION: Clone ID: 13392624.pep
US-10-767-701-53499

Query Match 14.2%; Score 110; DB 4; Length 123;
Best Local Similarity 29.7%; Pred. No. 0.22;
Matches 38; Conservative 7; Mismatches 27; Indels 56; Gaps 7;

QY 4 NDPAP-----GSDSAR--RPAP-----ALGPDPPAGWPFSGVPSRPICAASSAGLP 52
DB 28 NGPAPPPPLKGDSSRAVLTTPPPRGGLSPSSPSSG--SGPIPS----- 71
QY 53 PPVPPTMNDVTCSCGWSGCCIGPLISPMRWVVAAGNMPGVLEPGEGIPKIGFVV 112
DB 72 PXPXPP-----PLSSPSRPGPRARAPGSW-----PRLSA SPA 102
QY 113 LMTAPGSR 120
DB 103 RMTAAGRR 110

RESULT 6
US-09-919-039-278
Sequence 278, Application US/09919039
Publication No. US20030108871A1
GENERAL INFORMATION:
APPLICANT: Kaser, Matthew R.
TITLE OF INVENTION: GENES EXPRESSED IN TREATED HUMAN C3A LIVER CELL CULTURES
FILE REFERENCE: PA-0035 US
CURRENT APPLICATION NUMBER: US/09/919,039
CURRENT FILING DATE: 2002-09-09
PRIOR APPLICATION NUMBER: 60/222,113
PRIOR FILING DATE: 2000-07-28
NUMBER OF SEQ ID NOS: 401
SOFTWARE: PERL Program
SEQ ID NO 278
LENGTH: 707
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: miac_feature
OTHER INFORMATION: incyle ID No. US20030108871A1 2457215CD1
US-09-919-039-278

Query Match 13.4%; Score 103.5; DB 3; Length 707;
Best Local Similarity 28.0%; Pred. No. 4.2;
Matches 35; Conservative 6; Mismatches 41; Indels 43; Gaps 4;

QY 6 PPAAGSDSAR-----SRPAPALGDDPPASG-----WFDGIVPSRPICAA 45
DB 99 PPPPPQSSSKPVVAQGGPAPGVSSAPPASSAPPATPPISGAPGSGGPTPTPPPAVT 158
QY 46 SSSAGLPPVPPTMNDVTCSCGWSGCCIGPLISPMRWVVAAGNMPGVLEPGEGI 105

DB 159 SAPPAPPPPP-----SSGVPTTPPQAGGPPPPPAVPGGCP 196
QY 106 -PKIG 109
DB 197 GPXQG 201

RESULT 7
US-09-864-761-33360
Sequence 33360, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharron G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
FILE REFERENCE: Aesomica-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 33360
LENGTH: 135
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC006111.2
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 11
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 17
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 17
OTHER INFORMATION: EXPRESSED IN HEPA, SIGNAL = 20
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 11
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 12
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 15

OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 15
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 13
OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 18
OTHER INFORMATION: SWISSPROT HIT: Q01705, EVALUATE 9.30e-02
OTHER INFORMATION: EST_HUMAN HIT: A003023.1, EVALUATE 2.00e-06
US-09-864-761-33360

Query Match 13.2%; Score 102; DB 3; Length 135;
Best Local Similarity 26.3%; Pred. No. 1.1;
Matches 30; Conservative 8; Mismatches 46; Indels 30; Gaps 4;

QY 7 PAFGSDSARSRPAPALGPPASGWFDSGLVPSRPICAASSAGLPP-----PVP 57
DB 26 PGGWMASSRISGAQPSGRCRRTGTCSCASPSACIRGSMPSWPAPCPKASTRPTGP 85

QY 58 TWLNN-----DVTCCSGWVSCCIGPLISPSMPRVVA--AGSNMP 95
DB 86 RMLSSCRACAPMPQSVHGPSTCTCTACMCS-----TPSWPAAWRPRRAGPWP 134

RESULT 8
US-10-437-963-185237
Sequence 185237, Application US/10437963
Publication No. US20040123343A1
GENERAL INFORMATION:

APPLICANT: La Rosa, Thomas J.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
APPLICANT: Wu, Wei
APPLICANT: Boukharov, Andrey A.
APPLICANT: Barbazuk, Brad
APPLICANT: Li, Ping
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53221)B
CURRENT APPLICATION NUMBER: US/10/437,963
CURRENT FILING DATE: 2003-05-14
NUMBER OF SEQ ID NOS: 204966
SEQ ID NO 185237
LENGTH: 285
TYPE: PRT
ORGANISM: Oryza sativa
FEATURE:
NAME/KEY: unsure
LOCATION: (1)..(285)
OTHER INFORMATION: unsure at all Xaa locations
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT4530_82152C.1.pep
US-10-437-963-185237

Query Match 13.1%; Score 101.5; DB 4; Length 285;
Best Local Similarity 34.3%; Pred. No. 2.5;
Matches 34; Conservative 8; Mismatches 22; Indels 35; Gaps 8;

QY 6 PPA--PGSD-SARSRPAPALGPPD---PASGWFDSGLVPSRPICAASSAGLPPVPPTW 59
DB 70 PPAASPGTKGRRLSPASGPPSPPTPS-----PSP-ASASAPASPPWPPT- 119

QY 60 LNDVTCSCGWSGCCIGPLISPSMPRVVAAGMPTGV 98
DB 120 -----RGCC-----APTWRRTWPSAS---PSSV 139

RESULT 9
US-10-437-963-169481
Sequence 169481, Application US/10437963
Publication No. US20040123343A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei

APPLICANT: Wu, Wei
APPLICANT: Boukharov, Andrey A.
APPLICANT: Barbazuk, Brad
APPLICANT: Li, Ping
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53221)B
CURRENT APPLICATION NUMBER: US/10/437,963
CURRENT FILING DATE: 2003-05-14
NUMBER OF SEQ ID NOS: 204966
SEQ ID NO 169481
LENGTH: 179
TYPE: PRT
ORGANISM: Oryza sativa
FEATURE:
NAME/KEY: unsure
LOCATION: (1)..(179)
OTHER INFORMATION: unsure at all Xaa locations
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT4530_67899C.1.pep
US-10-437-963-169481

Query Match 12.9%; Score 99.5; DB 4; Length 179;
Best Local Similarity 28.0%; Pred. No. 2.3;
Matches 35; Conservative 8; Mismatches 37; Indels 45; Gaps 6;

QY 6 PPAP-----GSDSARS-----RPAPALGPPASGWFDSGLVPSRPICAASSAGL 51
DB 23 PPASAGEPLGCAESATAMITPPRPPELGPAPXRGPEPDG-----GL 69

QY 52 -----PPVPPT-----WLDVTCSCGWSGCCIGPLISPSMPRVVAAGSNMP 95
DB 70 GANCPPEPSPATRAVAVLLPMASTYAGRCG--VGHCLAPCBPPEPPASPAAGGWH 127

QY 96 TGVEL 100
DB 128 CGAGL 132

RESULT 10
US-10-029-386-33093
Sequence 33093, Application US/10029386
Publication No. US2003019704A1
GENERAL INFORMATION:

APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
FILE REFERENCE: AEOMICA-X-2
CURRENT APPLICATION NUMBER: US/10/029,386
CURRENT FILING DATE: 2001-12-20
NUMBER OF SEQ ID NOS: 34288
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 33093
LENGTH: 199
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: Map to AL049853.1
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 3.9
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3.2
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.6
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.9
OTHER INFORMATION: SWISSPROT HIT: P23206, EVALUATE 3.00e-08
US-10-029-386-33093

Query Match 12.9%; Score 99.5; DB 4; Length 199;
Best Local Similarity 26.6%; Pred. No. 2.5;
Matches 37; Conservative 15; Mismatches 56; Indels 31; Gaps 7;

QY 9 PGGSDSARSRPAP-----LSPD-----PPASGWFDSGLVPSRPICAA-----SSAGLPP 53
DB 120 -----LSPD-----PPASGWFDSGLVPSRPICAA-----SSAGLPP 53

Db 60 PGLAGTBMSPALAGWIGPPCCHLPPRAGMDHNLTCRPGLAGTTMPSPALGICP 119
Qy 54 P-----VPTWLN-----NDVTCSSGWS-----CCIGPLISPSWPRVWVAAGNWP 98
Db 120 PPGHLP--WAGMDHNLTCRPGLAGTTMPSPALAGWIGPPCCHLEQTRQERWGP 178
Qy 99 ELPEGIGPKIGFVVLMLAP 117
Db 179 KVPQTGAPOTWAGLMAHP 197

RESULT 11

US-09-925-302-741
; Sequence 741, Application US/09925302
; Patent No. US20020044941A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA104
; CURRENT APPLICATION NUMBER: US/09/925,302
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05918
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 896
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 741
; LENGTH: 212
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-302-741

Query Match 12.9%; Score 99.5; DB 3; Length 212;
Best Local Similarity 29.7%; Pred. No. 2.7;
Matches 43; Conservative 9; Mismatches 46; Indels 47; Gaps 9;

Qy 6 PPAPGSDSARS-----RPAPALGDDP-----ASGFDSGLVSPRPICA 44
Db 71 PPPPTQAKARAGLAVSAPSLTLPKTRLSSTSLANPLAPKXDPSPSPTTP-KR 129
Qy 45 ASSAGLPPVPPPTWLNNDVTCSSGWSCTIGPLISPSW--RVWVA-----AGNWP- 95
Db 130 GSPSSGLEPLVPS-----VCPRGPL--PRWPLGIRKAMALREGGRGMSG 174
Qy 96 TGVLPBGIGPKIGFVVLMLAPGSR 120
Db 175 CAIGVSGSFSARVG-VVWEGREASR 198

RESULT 12

US-09-925-302-741
; Sequence 741, Application US/09925302
; Publication No. US20030064072A9
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA104
; CURRENT APPLICATION NUMBER: US/09/925,302
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05918
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 896
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 741
; LENGTH: 212
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-302-741

Query Match 12.9%; Score 99.5; DB 3; Length 212;

Best Local Similarity 29.7%; Pred. No. 2.7;
Matches 43; Conservative 9; Mismatches 46; Indels 47; Gaps 9;

Qy 6 PPAPGSDSARS-----RPAPALGDDP-----ASGFDSGLVSPRPICA 44
Db 71 PPPPTQAKARAGLAVSAPSLTLPKTRLSSTSLANPLAPKXDPSPSPTTP-KR 129
Qy 45 ASSAGLPPVPPPTWLNNDVTCSSGWSCTIGPLISPSW--RVWVA-----AGNWP- 95
Db 130 GSPSSGLEPLVPS-----VCPRGPL--PRWPLGIRKAMALREGGRGMSG 174
Qy 96 TGVLPBGIGPKIGFVVLMLAPGSR 120
Db 175 CAIGVSGSFSARVG-VVWEGREASR 198

RESULT 13

US-10-437-963-163320
; Sequence 163320, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 163320
; LENGTH: 297
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_62327C.1.pep
US-10-437-963-163320

Query Match 12.9%; Score 99.5; DB 4; Length 297;
Best Local Similarity 25.1%; Pred. No. 3.8;
Matches 47; Conservative 13; Mismatches 50; Indels 77; Gaps 9;

Qy 2 RVNDPPAPGSDSARSAPAPA-----LGPDPASGFDSGLVPS-----R 40
Db 102 RRRPPAPA--PARSGAPAPAGCGCRLLPPPPPPAADSPLPHAPAAASACAASASIR 159
Qy 41 PICASSAGL-----PPVPP-----PTWL----- 60
Db 160 HMCSSSSAGICTPRFAMIDASFPATSSRRRAAPQTPRATPSPWAAAPSPRRRCS 219
Qy 61 ---NNDVTCSSGWSCTIGPLISPSWPRVWVAAGNWP-----TGVLPBGIGPKIGFVVL 113
Db 220 RRGRTAAAGKGR-----PRVAPRR--AAASSPQASSAGDIPTRGSPAAAGCVC 272
Qy 114 WLAPGSR 120
Db 273 GGGDGAR 279

RESULT 14

US-10-437-963-114935
; Sequence 114935, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei

```

1  APPLICANT: Mu, Wei
2  APPLICANT: Boukharov, Andrey A.
3  APPLICANT: Barbazuk, Brad
4  APPLICANT: Li, Ping
5  TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
6  TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
7  FILE REFERENCE: 38-21(53221) B
8  CURRENT APPLICATION NUMBER: US/10/437,963
9  CURRENT FILING DATE: 2003-05-14
10 NUMBER OF SEQ ID NOS: 204966
11 SEQ ID NO 114935
12 LENGTH: 327
13 TYPE: PRT
14 ORGANISM: Oryza sativa
15 FEATURES:
16 OTHER INFORMATION: Clone ID: PAT_MRT4530_18579C.1.pep
17 US-10-437-963-114935

```

QY	1	M R V N D P P A G S D S A R S P R P A L G P D P P A G M F D S G I V S R P I C A S S A G A L P P V P P M L	60
Db	1	M L P S P Q L P Q V T T A P A P S P V S T V S M - P P T S - - - - S A L L P S - - - - - P P P P L T C S	43
QY	61	N N D Y T C C S G M V S C C I G L P L S P S M F R M V A A G N M F T Q Y E L P E G S I P K I - G F V L M L A R G S	112
Db	44	R V V V A A A T A T S P A A G P L P P R - - - - - P A G C L P R A A I Y T L T P A P A P	85
QY	120	R I - - D A I G S F S K S V I T	134
Db	86	A V S S A R G S T F A K S T L T	102

```

RESULT 15
US-10-767-701-40674
: Sequence 40674, Application US/10767701
: Publication No. US20040172684A1
: GENERAL INFORMATION:
: APPLICANT: Kao, Yohsei
: APPLICANT: Zhou, Yihua
: TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
: TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
: FILE REFERENCE: 38-21(53535)B
: CURRENT FILING DATE: 2004-01-29
: NUMBER OF SEQ ID NOS: 63128
: SEQ ID NO 40674
: LENGTH: 155
: TYPE: PRT
: ORGANISM: Sorghum bicolor
: FEATURES:
: OTHER INFORMATION: Clone ID: SORBI-28MAY03-C80510_1.pep
: US-10-767-701-40674

```

Query Match 12.7%; Score 98; DB 4; Length 155;
Best Local Similarity 22.8%; Pred. No. 2.6;
Matches 31; Conservative 7; Mismatches 48; Indels 50; Gaps 3;

Qy	6	PPAAGSDASRSRPA	PALGDPDPPASGMFDSGLVSPSRICAASSAGLPPVPPTMNN-----	61
Dd	25	PPAPPPPPPAPPPLAPP	PPAPGARPRGRPPAA--PAGRAAAPAVPVVRPMWGGRCRG	82
Qy	62	-----	-NDVTCCSGWSCCGIPLISPSWP	84
Dd	83	PGAAPPPPSRAA	PTRRPRAAPPAPTARGRCCGVSPSCCORSALSCSPSPGFFWAMP	144
Qy	85	RVMVAAGNMPTGTVEL	100	
Dd	143	R-----RWPTSVPL	151	

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocelebration Ltd.

OM protein - protein search, using sw model

Run on: April 14, 2006, 18:37:38 ; Search time 5.95625 Seconds
(without alignments)
991.565 Million cell updates/sec

Title: US-10-620-246-90

Perfect score: 774
Sequence: 1 MRVNDPPAPGSDARSRRPAP.....RIDAGSSFYSKXVLTAVSAW 139

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 217505 seqs, 42489236 residues

Total number of hits satisfying chosen parameters: 217505

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_AA_New:
1: /SIDS5/ptodata/2/pubpaa/US08_NEW_PUB pep.*
2: /SIDS5/ptodata/2/pubpaa/US06_NEW_PUB pep.*
3: /SIDS5/ptodata/2/pubpaa/US07_NEW_PUB pep.*
4: /SIDS5/ptodata/2/pubpaa/PCT_NEW_PUB pep.*
5: /SIDS5/ptodata/2/pubpaa/US05_NEW_PUB pep.*
6: /SIDS5/ptodata/2/pubpaa/US10_NEW_PUB pep.*
7: /SIDS5/ptodata/2/pubpaa/US11_NEW_PUB pep.*
8: /SIDS5/ptodata/2/pubpaa/US60_NEW_PUB pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	101.5	13.1	584	US-11-051-720-324	Sequence 324, App
2	101.5	13.1	584	US-11-043-806-368	Sequence 368, App
3	101	13.0	481	US-11-087-099-11981	Sequence 11981, A
4	96.5	12.5	1049	US-10-979-095-7	Sequence 7, Appl
5	91.5	11.8	398	US-11-150-845-10	Sequence 10, Appl
6	91.5	11.8	398	US-11-150-847-10	Sequence 10, Appl
7	91.5	11.8	502	US-11-122-795-14	Sequence 14, Appl
8	91.5	11.8	502	US-11-150-845-2	Sequence 2, Appl
9	91.5	11.8	502	US-11-149-945-1	Sequence 1, Appl
10	91.5	11.8	502	US-11-150-847-2	Sequence 2, Appl
11	91.5	11.8	502	US-11-079-907-28	Sequence 28, Appl
12	91.5	11.8	609	US-11-150-845-20	Sequence 20, Appl
13	91.5	11.8	609	US-11-150-847-20	Sequence 20, Appl
14	91.5	11.8	647	US-11-087-099-7887	Sequence 7887, App
15	91.5	11.8	647	US-11-188-298-18316	Sequence 18316, A
16	91.5	11.8	649	US-11-150-845-18	Sequence 18, Appl
17	91.5	11.8	649	US-11-150-845-22	Sequence 22, Appl
18	91.5	11.8	649	US-11-150-847-18	Sequence 18, Appl
19	91.5	11.8	649	US-11-150-847-22	Sequence 22, Appl
20	91.5	11.8	715	US-11-150-845-14	Sequence 14, Appl
21	91.5	11.8	715	US-11-150-847-14	Sequence 14, Appl
22	90.5	11.7	408	US-11-150-845-12	Sequence 12, Appl
23	90.5	11.7	408	US-11-150-847-12	Sequence 12, Appl
24	90.5	11.7	497	US-11-087-099-809	Sequence 809, App
25	90.5	11.7	505	US-11-134-563-6	Sequence 6, Appl

26	90.5	11.7	505	7	US-11-150-845-4	Sequence 4, Appl
27	90.5	11.7	505	7	US-11-149-945-2	Sequence 2, Appl
28	90.5	11.7	505	7	US-11-150-847-4	Sequence 4, Appl
29	90.5	11.7	638	7	US-11-150-845-24	Sequence 24, Appl
30	90.5	11.7	638	7	US-11-150-847-24	Sequence 24, Appl
31	88	11.4	667	7	US-11-087-099-9879	Sequence 9879, Ap
32	87.5	11.3	277	7	US-11-096-568A-1871	Sequence 1871, Ap
33	87.5	11.3	277	7	US-11-096-568A-27199	Sequence 27199, A
34	87	11.2	412	7	US-11-096-568A-21231	Sequence 21231, A
35	87	11.2	533	7	US-11-124-368A-335	Sequence 335, App
36	86	11.1	229	7	US-11-072-512-3048	Sequence 3048, Ap
37	86	11.1	315	7	US-11-072-512-3207	Sequence 3207, Ap
38	85.5	11.0	716	7	US-11-150-845-16	Sequence 16, Appl
39	85.5	11.0	716	7	US-11-150-847-16	Sequence 16, Appl
40	83	10.7	137	6	US-10-644-807-349	Sequence 349, App
41	82.5	10.7	283	7	US-11-096-568A-21223	Sequence 21223, A
42	82.5	10.7	761	6	US-10-204-639-19	Sequence 19, Appl
43	82	10.6	1870	7	US-11-188-298-17645	Sequence 17645, A
44	81.5	10.5	319	7	US-11-096-568A-24535	Sequence 24535, A
45	81	10.5	273	7	US-11-096-568A-22109	Sequence 22109, A

ALIGNMENTS

```
RESULT 1
US-11-051-720-324
; Sequence 324, Application US//11051720
; Publication No. US20060046257A1
; GENERAL INFORMATION:
; APPLICANT: Compugen Ltd
; TITLE OF INVENTION: NOVEL NUCLEOTIDE AND AMINO ACID SEQUENCES, AND ASSAYS AND METHOD
; FILE REFERENCE: 1847.1002
; CURRENT APPLICATION NUMBER: US/11/051.720
; CURRENT FILING DATE: 2005-01-27
; NUMBER OF SEQ ID NOS: 1780
; SEQ ID NO 324
; LENGTH: 584
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-051-720-324

Query Match      13.1%, Score 101.5; DB 7; Length 584;
Best Local Similarity 26.3%; Pred. No. 0.13;
Matches 41; Conservative 16; Mismatches 58; Indels 41; Gaps 7;

QY 18 PAPALGPPAPAGFDSGLVPSRPICAASS-AGLPPEVPTMINNDVTCSSGWSCTIG 76
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 371 PDPPGCGVPTGW--GPTLPSRAVLARSSAGCGPRTVSTAPGMCIGCSFG--LCVG 425
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

QY 77 -----PLISFWRWVVAAGN-----WPTGVLPGRGIRKIGPVVLMIA----- 116
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 426 VPLPTSPPLADADKPPAPVASSSATSLPWPVVIIGIPAGAVITLGLMLLCOAQKCP 485
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

QY 117 -----PGSRIDAIGSSFYSKXVLTAVSA 138
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 486 CTPAPAPLPCHRRPRTGARDRSGDKLP-LLALISA 520
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

RESULT 2
US-11-043-806-368
; Sequence 368, Application US//11043806
; Publication No. US20060051774A1
; GENERAL INFORMATION:
; APPLICANT: Compugen Ltd
; TITLE OF INVENTION: Novel Nucleotide and Amino Acid Sequences, and Assays and Method
; FILE REFERENCE: 1847.1003
; CURRENT APPLICATION NUMBER: US/11/043.806
; CURRENT FILING DATE: 2005-01-27
; NUMBER OF SEQ ID NOS: 575
; SEQ ID NO 368
```

```
/ LENGTH: 584
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-11-043-806-368

Query Match          13.1%; Score 101.5; DB 7; Length 584;
Best Local Similarity 26.3%; Pred. No. 0.13;
Matches 41; Conservative 16; Mismatches 58; Indels 41; Gaps 7;

18 PAPALPDPAPSGWFDGVLPSRPRICASS--AGLPVPVPTMLNNDVTCGSGWVSCCIG 76
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 371 PDPPGGVPTG--GTLPSRAVLARSSAEGQPGTGTITAPGGLGCSPG---LCVG 425
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

77 -----PLISPSWPRVWVAAGN-----WPTGVLPSEGIPKIGFVVLMLA----- 116
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 426 VPLPTEPRLALADPKPPGPVVAASSSATSLPMVPVIGIPAGAVPILGTLLMLTCAQAKP 485
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

117 -----PGSRIDALGSSFSKSVLTAVSA 138
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 486 CTAPAPPLPGHRPPTGARDRSGDKDLP--LAALSA 520
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 3
US-11-087-099-11981
/ Sequence 11981, Application US/11087099
/ Publication No. US20060041961A1
/ GENERAL INFORMATION:
/ APPLICANT: Abad, Mark S. et al.
/ TITLE OF INVENTION: Genes and Uses for Plant Improvement
/ FILE REFERENCE: 38-21(53450)B EP
/ CURRENT APPLICATION NUMBER: US/11/087,099
/ CURRENT FILING DATE: 2005-03-22
/ NUMBER OF SEQ ID NOS: 12464
/ SEQ ID NO 11981
/ LENGTH: 481
/ TYPE: PRT
/ ORGANISM: Triticum aestivum
US-11-087-099-11981

Query Match          13.0%; Score 101; DB 7; Length 481;
Best Local Similarity 28.8%; Pred. No. 0.12;
Matches 38; Conservative 7; Mismatches 41; Indels 46; Gaps 5;

6 PP---APGSDASRPAALGDPDPASGWFDSGLVPSRPRICASSAGLPPVPPTMLAN 62
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 8 PPGVEGCAATYRCRGYAPRPPSPG---KLLPPPALAHPPALAPPPPPPLV-- 61
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

63 DVTCCSGWVSCCIGPLISPSWPRVWVAAGNWPVTGVELPBGGLPKIGFVVLMLAPGSRID 122
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 62 --VCGGGSVAPCSG-----GRIIP-----PPSPGT 84
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

123 AIGSSPSKSVLT 134
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 85 ALGPTNSKSTFT 96
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 4
US-10-979-095-7
/ Sequence 7, Application US/10979095
/ Publication No. US20060068481A1
/ GENERAL INFORMATION:
/ APPLICANT: INCYTE GENOMICS, INC.
/ APPLICANT: YANG, Junming
/ APPLICANT: BAUGHN, Mariah R.
/ APPLICANT: BURFORD, Neil
/ APPLICANT: AU-YOUNG, Janice
/ APPLICANT: LU, Dying Alina M.
/ APPLICANT: REDDY, Roopa
/ APPLICANT: YUE, Henry
/ APPLICANT: YAO, Monique G.
/ APPLICANT: LAU, Preeti
/ APPLICANT: KAHN, Farrah A.
/ TITLE OF INVENTION: HUMAN KINASES
```

```
/ FILE REFERENCE: PI-0002 PCT
/ CURRENT APPLICATION NUMBER: US/10/979,095
/ CURRENT FILING DATE: 2004-11-02
/ PRIOR APPLICATION NUMBER: US/10/168,582
/ PRIOR FILING DATE: 2002-06-20
/ PRIOR APPLICATION NUMBER: 60/172,066; 60/176,107; 60/176,107; 60/177,731
/ PRIOR FILING DATE: 1999-12-23; 2000-01-14; 2000-01-14; 2000-01-21
/ NUMBER OF SEQ ID NOS: 24
/ SOFTWARE: PERL Program
/ SEQ ID NO 7
/ LENGTH: 1049
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: misc feature
/ OTHER INFORMATION: Incyte ID No: 3210906CD1
US-10-979-095-7

Query Match          12.5%; Score 96.5; DB 6; Length 1049;
Best Local Similarity 27.3%; Pred. No. 0.64;
Matches 38; Conservative 10; Mismatches 46; Indels 45; Gaps 8;

9 PGSDASRPAALGDPDPASGWFDSGLVPSRPI--CASSAGLPP--VPTMLNNDV 64
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 924 PAAALAGTPAP-----PPAPAP-----PSRPVPSGAHWSHGPPPCMPAPAKQPSL 972
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

65 TCCSG--WVSCCIGPLISPSWPRVW-----VAAGNWPVTGVELPBGGLPK 107
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 973 LAPPGPNN-----LGPPTGSTRGALLLRNSPQLRLRAASCGSGSENVGPAAVAP- 1027
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

108 IGFVVLMLAPGSRIDALIGS 126
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1028 -----GPLSRSTSVAS 1038
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 5
US-11-150-845-10
/ Sequence 10, Application US/11150845
/ Publication No. US20060003399A1
/ GENERAL INFORMATION:
/ APPLICANT: Cytokinetics, Inc.
/ APPLICANT: Tomasevic, Nenad
/ APPLICANT: Jia, Zhiheng
/ APPLICANT: Sakowicz, Roman
/ APPLICANT: Pierce, Daniel
/ APPLICANT: Finer, Jeffrey
/ TITLE OF INVENTION: HIGH THROUGHPUT ACTIN POLYMERIZATION ASSAY
/ FILE REFERENCE: 020552-007720US
/ CURRENT APPLICATION NUMBER: US/11/150,845
/ CURRENT FILING DATE: 2005-06-10
/ PRIOR APPLICATION NUMBER: US 60/673,444
/ PRIOR FILING DATE: 2005-04-20
/ PRIOR APPLICATION NUMBER: 60/578,949
/ PRIOR FILING DATE: 2004-06-10
/ NUMBER OF SEQ ID NOS: 78
/ SOFTWARE: PatentIn version 3.3
/ SEQ ID NO 10
/ LENGTH: 398
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (1)-(398)
/ OTHER INFORMATION: 105WASP
US-11-150-845-10

Query Match          11.8%; Score 91.5; DB 7; Length 398;
Best Local Similarity 27.9%; Pred. No. 0.68;
Matches 36; Conservative 5; Mismatches 43; Indels 45; Gaps 5;

6 PPAPGSDASRPAALGDPDPASGWFDSGLVPSRPRICASSAGLPP 53
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 225 PPIVGKNGKSGPLVPVIGIADPPPTPRGPPPPGKG-----PPPPPPATGSGPLP 279
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```


RESULT 9

US-11-149-945-1
Sequence 1, Application US/11149945
Publication No. US20060014266A1
GENERAL INFORMATION:
APPLICANT: Cycokinetics, Inc.
APPLICANT: Tomasevic, Nenad
APPLICANT: Sakowicz, Roman
TITLE OF INVENTION: PURIFICATION OF ARP2/3 COMPLEX AND COMPOSITIONS CONTAINING
FILE REFERENCE: 020552-007810US
CURRENT FILING DATE: 2005-06-10
PRIOR APPLICATION NUMBER: US 60/578,969
PRIOR FILING DATE: 2004-06-10
NUMBER OF SEQ ID NOS: 11
SOFTWARE: PatentIn version 3.3
SEQ ID NO 1
LENGTH: 502
TYPE: PRT
ORGANISM: Homo sapiens
US-11-149-945-1

Query Match 11.8%; Score 91.5; DB 7; Length 502;
Best Local Similarity 27.9%; Pred. No. 0.85;
Matches 36; Conservative 5; Mismatches 43; Indels 45; Gaps 5;

QY 6 PPAAGSDSARSAPAL-----GPPPASGWFDSGLVPSRPICASSSAGLPP 53
DB 329 PPIVGNKGRSGPLPVPVLGIAPPPTPRGPPPPRGG-----PPPPPPATGRSGPLPP 383
QY 54 PVP-----PTWLNDVTCSSGWSGCCIGPLISPSWPRVWVAAGNWPVGVELPGEIGPKIG 109
DB 384 PPGAGGP-----PMPPPPPPPPSSGNGAPAPPLPALVPAG 424
QY 110 FVVLMLAPG 118
DB 425 -----LAPG 428

RESULT 10

US-11-150-487-2
Sequence 2, Application US/11150487
Publication No. US20060024786A1
GENERAL INFORMATION:
APPLICANT: Cycokinetics, Inc.
APPLICANT: Tomasevic, Nenad
APPLICANT: Russell, Alan
APPLICANT: Wang, Manning
APPLICANT: Sakowicz, Roman
APPLICANT: Betaud, Christophe
TITLE OF INVENTION: WASP AND N-WASP CONSTRUCTS AND METHODS OF EXPRESSING SUCH
FILE REFERENCE: 020552-007910US
CURRENT FILING DATE: 2005-06-10
PRIOR APPLICATION NUMBER: US/11/150,487
PRIOR FILING DATE: 2004-06-10
NUMBER OF SEQ ID NOS: 70
SOFTWARE: PatentIn version 3.3
SEQ ID NO 2
LENGTH: 502
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)..(502)
OTHER INFORMATION: FL-WASP
US-11-150-487-2

Query Match 11.8%; Score 91.5; DB 7; Length 502;
Best Local Similarity 27.9%; Pred. No. 0.85;

Matches 36; Conservative 5; Mismatches 43; Indels 45; Gaps 5;

QY 6 PPAAGSDSARSAPAL-----GPPPASGWFDSGLVPSRPICASSSAGLPP 53
DB 329 PPIVGNKGRSGPLPVPVLGIAPPPTPRGPPPPRGG-----PPPPPPATGRSGPLPP 383
QY 54 PVP-----PTWLNDVTCSSGWSGCCIGPLISPSWPRVWVAAGNWPVGVELPGEIGPKIG 109
DB 384 PPGAGGP-----PMPPPPPPPPSSGNGAPAPPLPALVPAG 424
QY 110 FVVLMLAPG 118
DB 425 -----LAPG 428

RESULT 11

US-11-079-907-28
Sequence 28, Application US/11079907
Publication No. US20060029946A1
GENERAL INFORMATION:
APPLICANT: Hahn, Klaus M.
TITLE OF INVENTION: Live Cell Biosensors
FILE REFERENCE: 1361.044US1
CURRENT FILING DATE: 2005-03-14
PRIOR APPLICATION NUMBER: US 60/552,663
PRIOR FILING DATE: 2004-03-12
NUMBER OF SEQ ID NOS: 28
SOFTWARE: PasteSeq for Windows Version 4.0
SEQ ID NO 28
LENGTH: 502
TYPE: PRT
ORGANISM: Homo sapiens
US-11-079-907-28

Query Match 11.8%; Score 91.5; DB 7; Length 502;
Best Local Similarity 27.9%; Pred. No. 0.85;
Matches 36; Conservative 5; Mismatches 43; Indels 45; Gaps 5;

QY 6 PPAAGSDSARSAPAL-----GPPPASGWFDSGLVPSRPICASSSAGLPP 53
DB 329 PPIVGNKGRSGPLPVPVLGIAPPPTPRGPPPPRGG-----PPPPPPATGRSGPLPP 383
QY 54 PVP-----PTWLNDVTCSSGWSGCCIGPLISPSWPRVWVAAGNWPVGVELPGEIGPKIG 109
DB 384 PPGAGGP-----PMPPPPPPPPSSGNGAPAPPLPALVPAG 424
QY 110 FVVLMLAPG 118
DB 425 -----LAPG 428

RESULT 12

US-11-150-845-20
Sequence 20, Application US/11150845
Publication No. US20060003399A1
GENERAL INFORMATION:
APPLICANT: Cycokinetics, Inc.
APPLICANT: Tomasevic, Nenad
APPLICANT: Jia, Zhiheng
APPLICANT: Sakowicz, Roman
APPLICANT: Pierce, Daniel
APPLICANT: Finer, Jeffrey
TITLE OF INVENTION: HIGH THROUGHPUT ACTIN POLYMERIZATION ASSAY
FILE REFERENCE: 020552-007720US
CURRENT FILING DATE: 2005-06-10
PRIOR APPLICATION NUMBER: US 60/673,444
PRIOR FILING DATE: 2005-04-20
PRIOR APPLICATION NUMBER: 60/578,949
PRIOR FILING DATE: 2004-06-10
NUMBER OF SEQ ID NOS: 78

```
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 20
; LENGTH: 609
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Myc-105WASP-TAP
US-11-150-845-20
```

```
Query Match          11.8%; Score 91.5; DB 7; Length 609;
Best Local Similarity 27.9%; Pred. No. 1;
Matches 36; Conservative 5; Mismatches 43; Indels 45; Gaps 5;
```

```
Qy 6 PPAGSDSARSAPAL-----GDP PASGWFDSGLVPSRPTCAASSAGLPP 53
Db 247 PPVIGKMGKRGSLPPVPLGIAPPPTRPGPPRGSG-----PPPPPPATGRSGPLP 301
Qy 54 PVP-----PTWLNNDVTCSCGWSGCCIGPLISPSWPRVVAAGWPTGVLEPGSGIKTG 109
Db 302 PPPGAGGP-----FMPPPPPPPPSSGNGGAPAPPLPPALVPAAG 342
Qy 110 FVVVLMAPG 118
Db 343 -----LAPG 346
```

RESULT 13

```
US-11-150-487-20
; Sequence 20, Application US/11150487
; Publication No. US20060024786A1
; GENERAL INFORMATION:
; APPLICANT: Cytokinetics, Inc.
; APPLICANT: Tomasevic, Nenad
; APPLICANT: Russell, Alan
; APPLICANT: Wang, Maming
; APPLICANT: Sakowicz, Roman
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: WASP AND N-WASP CONSTRUCTS AND METHODS OF EXPRESSING SUCH
; TITLE OF INVENTION: CONSTRUCTS
; FILE REFERENCE: 020552-007910US
; CURRENT APPLICATION NUMBER: US/11/150,487
; CURRENT FILING DATE: 2005-06-10
; PRIOR APPLICATION NUMBER: US 60/578,913
; PRIOR FILING DATE: 2004-06-10
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 20
; LENGTH: 609
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Myc-105WASP-TAP
US-11-150-487-20
```

```
Query Match          11.8%; Score 91.5; DB 7; Length 609;
Best Local Similarity 27.9%; Pred. No. 1;
Matches 36; Conservative 5; Mismatches 43; Indels 45; Gaps 5;
```

```
Qy 6 PPAGSDSARSAPAL-----GDP PASGWFDSGLVPSRPTCAASSAGLPP 53
Db 247 PPVIGKMGKRGSLPPVPLGIAPPPTRPGPPRGSG-----PPPPPPATGRSGPLP 301
Qy 54 PVP-----PTWLNNDVTCSCGWSGCCIGPLISPSWPRVVAAGWPTGVLEPGSGIKTG 109
Db 302 PPPGAGGP-----FMPPPPPPPPSSGNGGAPAPPLPPALVPAAG 342
Qy 110 FVVVLMAPG 118
Db 343 -----LAPG 346
```

```
RESULT 14
US-11-087-099-7887
```

```
; Sequence 7887, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 7887
; LENGTH: 647
; TYPE: PRT
; ORGANISM: Brassica napus
US-11-087-099-7887
```

```
Query Match          11.8%; Score 91.5; DB 7; Length 647;
Best Local Similarity 40.0%; Pred. No. 1.1;
Matches 22; Conservative 9; Mismatches 21; Indels 3; Gaps 2;
```

```
Qy 6 PPAGSDSARSAPALGDP--PASGWFDSGLVPSRPTCAASSAGLPPVPPT 58
Db 12 PPSPPSNTTTTTPPASAPPTTSSPPPS--TFTSPPSRSTPSAPPPSPPT 65
```

RESULT 15

```
US-11-188-298-18316
; Sequence 18316, Application US/11188298
; Publication No. US20060075522A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: GENES AND USES FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-21(53452)B
; CURRENT APPLICATION NUMBER: US/11/188,298
; CURRENT FILING DATE: 2005-07-22
; PRIOR APPLICATION NUMBER: 60/592,978
; PRIOR FILING DATE: 2004-07-31
; NUMBER OF SEQ ID NOS: 22569
; SEQ ID NO 18316
; LENGTH: 647
; TYPE: PRT
; ORGANISM: Brassica napus
US-11-188-298-18316
```

```
Query Match          11.8%; Score 91.5; DB 7; Length 647;
Best Local Similarity 40.0%; Pred. No. 1.1;
Matches 22; Conservative 9; Mismatches 21; Indels 3; Gaps 2;
```

```
Qy 6 PPAGSDSARSAPALGDP--PASGWFDSGLVPSRPTCAASSAGLPPVPPT 58
Db 12 PPSPPSNTTTTTPPASAPPTTSSPPPS--TFTSPPSRSTPSAPPPSPPT 65
```

```
Search completed: April 14, 2006, 18:42:16
Job time : 6.95625 secs
```

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using SW model

Run on: April 14, 2006, 17:03:40 ; Search time 25.9599 Seconds
(without alignments)
1133.996 Million cell updates/sec

Title: US-10-620-246-147

Perfect score: 359
Sequence: 1 MPQSTVWFMNAKGFPIAP.....VEFEIGHSPKQPNATGVRSI 67

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_21.*
1: geneseqp1980s.*
2: geneseqp1990s.*
3: geneseqp2000s.*
4: geneseqp2001s.*
5: geneseqp2002s.*
6: geneseqp2003as.*
7: geneseqp2003bs.*
8: geneseqp2004s.*
9: geneseqp2005s.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	359	100.0	67	2	AAW72934
2	359	100.0	67	2	AAW72934
3	282	78.6	67	4	AAW78912
4	282	78.6	67	4	AAW78911
5	282	78.6	67	4	AAW78911
6	282	78.6	67	9	ADZ21478
7	275	76.6	67	4	AAW78914
8	275	76.6	67	4	AAW78914
9	256	71.3	67	6	ABU3189
10	227.5	63.4	88	7	ADH86286
11	223	62.1	70	7	ABM78962
12	223	62.1	70	9	ADZ21446
13	222	61.8	131	7	ABO66959
14	221	61.6	70	7	ADH54477
15	218.5	60.9	66	5	ABW07338
16	217	60.4	70	7	ADH54476
17	216	60.2	67	9	ABM93150
18	216	60.2	68	8	ADP91418
19	216	60.2	69	2	AAW46470
20	216	60.2	69	5	ABW07333
21	216	60.2	69	5	ABW07336
22	216	60.2	69	9	ADZ21440
23	216	60.2	73	7	ABO67609
24	214	59.6	73	7	ADP04023

25	213.5	59.5	66	6	ABU29407	Abu29407 Protein e
26	213.5	59.5	69	5	ABP27075	ABP27075 Streptoco
27	213	59.3	67	9	ADW11153	ADW11153 Cold shoc
28	212.5	59.2	73	7	ADC96608	ADC96608 E. faeciu
29	211.5	58.9	67	5	ABP27074	ABP27074 Streptoco
30	211.5	58.9	67	8	ADV89793	ADV89793 Streptoco
31	211.5	58.9	67	8	ADV81046	ADV81046 Streptoco
32	211.5	58.9	67	8	ADV83185	ADV83185 Streptoco
33	211	58.8	68	8	ADF91417	ADF91417 E. coli c
34	211	58.8	70	2	AAW06609	AAW06609 Cold shoc
35	211	58.8	70	2	AAW46472	AAW46472 Amino aci
36	211	58.8	70	2	AAW06665	AAW06665 E.coli co
37	211	58.8	70	3	AAW15971	AAW15971 E. coli p
38	211	58.8	70	4	AAE08868	AAE08868 Escherich
39	211	58.8	70	5	ABW07346	ABW07346 Cold-shoc
40	211	58.8	70	5	ABW07334	ABW07334 Cold-shoc
41	211	58.8	70	5	AAU77662	AAU77662 B. coli c
42	211	58.8	70	6	ABU31901	ABU31901 Protein e
43	211	58.8	70	6	ABU47382	ABU47382 Protein e
44	211	58.8	70	6	ABU14715	ABU14715 Protein e
45	211	58.8	70	6	ABU47535	ABU47535 Protein e

ALIGNMENTS

RESULT 1	
AAW72934	AAW72934 standard; protein; 67 AA.
ID	AAW72934
AC	AAW72934;
DT	21-JAN-1999 (first entry)
XX	
DB	Mycobacterium tuberculosis antigen CFP7B.
XX	
KW	Mycobacterium tuberculosis; antigen; vaccine; immunological; immunogen;
KW	infection.
XX	
OS	Mycobacterium tuberculosis.
XX	
PN	W09844119-A1.
XX	
PD	08-OCT-1998.
XX	
PF	01-APR-1998; 98MO-DK000132.
XX	
PR	02-APR-1997; 97DK-00000376.
PR	18-APR-1997; 97US-0044624P.
PR	10-NOV-1997; 97DK-00001277.
PR	05-JAN-1998; 98US-0070488P.
XX	
PA	(STAT-) STATENS SERUM INST.
XX	
PI	Andersen P, Nielsen R, Rosenkrands I, Weldingh K, Rasmussen PB,
PI	Oettinger T, Florio W;
DR	NPI, 1998-542705/46.
DR	N-PSDB; AAV63944.
XX	
PT	New isolated mycobacteria polypeptides and nucleic acids - used for
PT	developing products for the diagnosis of or vaccination against
PT	mycobacterial infections, particularly tuberculosis.
XX	
XX	Claim 1; Page 222; 163pp; English.
PS	
XX	
CC	The present sequence represents a Mycobacterium tuberculosis protein.
CC	Products from the present invention, which describes protein fragments
CC	and nucleic acid fragments derived from M. tuberculosis, can be used in
CC	the detection of and prevention of mycobacterial infections. In
CC	particular, the proteins and nucleic acids can be used for the diagnosis
CC	of or vaccination against tuberculosis caused by M. tuberculosis, M.
CC	africanum or M. bovis

XX Sequence 67 AA;
SQ
Query Match 100.0%; Score 359; DB 2; Length 67;
Best Local Similarity 100.0%; Pred. No. 2.4e-36;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MPOGTVKMFNAEKGFPIAPEDGSADV FVHYTEIQTGFRITLBNQKVEFEIGHSPKGPQ 60
DB 1 MPOGTVKMFNAEKGFPIAPEDGSADV FVHYTEIQTGFRITLBNQKVEFEIGHSPKGPQ 60
QY 61 ATGVRSI 67
DB 61 ATGVRSI 67
RESULT 2
AAV21954
ID AAV21954 standard; protein; 67 AA.
AC AAV21954;
XX
XX
DT 06-SRP-1999 (first entry)
XX
XX Amino acid sequence of antigen CFP7B.
DE
XX Immunogenic; Mycobacterium tuberculosis; immune response; infection;
KM tuberculosis; fusion polypeptide; T-cell epitope; ESAT-6; MPT59; TB;
KM pharmaceutical; vaccination; M. africanum; M. bovis; CFP7A; CFP30A;
KM CFP7B; CFP19; CFP27; CFP30A; RD1-ORF; CFP10A; CFP16; CFP19; CFP23;
KM CFP25A; CFP30B; CFP7B.
XX
XX Mycobacterium tuberculosis.
OS
XX
XX WO924577-A1.
PN
XX 20-MAY-1999.
PD
XX
XX 08-OCT-1998; 98WO-DK00438.
PF
XX
XX 10-NOV-1997; 97DK-00001277.
PR 05-JAN-1998; 98US-0070488P.
PR 01-APR-1998; 98WO-DK000132.
XX
XX (STAT-) STATENS SERUM INST.
PA
XX
XX Andersen P, Skjot R;
PI
XX
XX WPI; 1999-347282/29.
DR N-PSDB; AAX81093.
XX
XX
XX New immunogenic fragment of Mycobacterium tuberculosis.
PT
XX
XX Example 3; Page 81; 265pp; English.
PS
XX
XX The invention describes a substantially pure immunogenic polypeptide
CC fragment (I) from Mycobacterium tuberculosis that is able to evoke a
CC protective immune response against infections by mycobacteria belonging
CC to the tuberculosis complex. The invention provides a (1) fusion
CC polypeptide comprising at least one polypeptide fragment (I) and at least
CC one fusion partner; (2) a fusion polypeptide fragment comprising a T-cell
CC epitope from M. tuberculosis protein ESAT-6, or MPT59 and a second
CC different amino acid sequence from M. tuberculosis, and/or including a
CC sequence which protects the first amino acid sequence from in vivo
CC degradation or post-translational processing; (3) a nucleic acid fragment
CC that encodes the above polypeptides. The polypeptides and nucleic acid
CC are useful as pharmaceuticals, for diagnosis of and as antigens for
CC vaccination against TB caused by Mycobacterium tuberculosis, africanum or
CC bovis. The polypeptides are also useful for diagnosing ongoing or
CC previous sensitization in an animal with bacteria belonging to the
CC tuberculosis complex. The invention also describes the use of CFP7A or
CC CFP30A or a T-cell epitope of for the induction of a strong immune
CC response in a mammal; use of CFP7B, CFP19 or MPT59-ESAT6 or a T-cell

CC epitope of for diagnosis of TB in a mammal by performing a DTH type skin
CC test; use of CFP27, CFP30A, RD1-ORF2, RD1-ORF3, RD1-ORF5, MPT59-ESAT6,
CC ESAT6-MPT59, CFP10A, CFP19, CFP23, CFP25A, CFP30B, CFP7B or a T-
CC cell epitope of for the preparation of an immunological composition; and
CC for the preparation of a subunit vaccine
XX
XX Sequence 67 AA;
SQ
Query Match 100.0%; Score 359; DB 2; Length 67;
Best Local Similarity 100.0%; Pred. No. 2.4e-36;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MPOGTVKMFNAEKGFPIAPEDGSADV FVHYTEIQTGFRITLBNQKVEFEIGHSPKGPQ 60
DB 1 MPOGTVKMFNAEKGFPIAPEDGSADV FVHYTEIQTGFRITLBNQKVEFEIGHSPKGPQ 60
QY 61 ATGVRSI 67
DB 61 ATGVRSI 67
RESULT 3
AAB78912
ID AAB78912 standard; protein; 67 AA.
AC AAB78912;
XX
XX
DT 30-APR-2001 (first entry)
XX
XX C. glutamicum SRT protein sequence SEQ ID NO:84.
DE
XX Corynebacterium glutamicum; stress; resistance; tolerance; SRT;
KM fine chemical production; organic acid; proteinogenic amino acid;
KM nonproteinogenic amino acid; purine base; pyrimidine base; nucleoside;
KM nucleotide; lipid; saturated fatty acid; unsaturated fatty acid; diol;
KM carbohydrate; aromatic compound; vitamin; cofactor; polyketide; enzyme;
KM evolutionary study; environmental hazard; fermentation.
XX
XX
XX Corynebacterium glutamicum.
OS
XX
XX WO200100804-A2.
PN
XX
XX 04-JAN-2001.
PD
XX
XX 23-JUN-2000; 2000WO-IB000922.
PF
XX
XX 25-JUN-1999; 99US-0141031P.
PR 01-JUL-1999; 99DE-0103042P.
PR 01-JUL-1999; 99US-0142892P.
PR 08-JUL-1999; 99DE-01031413.
PR 08-JUL-1999; 99DE-01031457.
PR 08-JUL-1999; 99DE-01031541.
PR 09-JUL-1999; 99DE-01032209.
PR 09-JUL-1999; 99DE-01032230.
PR 14-JUL-1999; 99DE-01032914.
PR 27-AUG-1999; 99DE-01040764.
PR 27-AUG-1999; 99US-0151214P.
PR 31-AUG-1999; 99DE-01041382.
XX
XX (BADT) BASF AG.
PA
XX
XX Pompejus M, Kroegeer B, Schroeder H, Zeidler O, Haberhauer G;
PI
XX
XX Lee H, Kim H;
DR WPI; 2001-061972/07.
DR N-PSDB; AAF71025.
XX
XX New isolated Corynebacterium glutamicum nucleic acid encoding a stress,
PT tolerance or resistance protein, for production or modulation of
PT production of fine chemicals, such as, e.g. amino acids, lipids,
PT carbohydrates, or enzymes.
XX
XX Claim 20; Page 231-232; 526pp; English.

XX AAF70984 to AAF71133 encode the *Corynebacterium glutamicum* stress,
 CC resistance, and tolerance (SRT) proteins given in AAB78871 to AAB79020.
 CC The C. glutamicum SRT genes (I) can be used in vectors (II) for
 CC expression in host cells and production of fine chemicals, such as, an
 CC organic acid, a proteinogenic or nonproteinogenic amino acid (preferred),
 CC a purine or pyrimidine base, a nucleoside, a nucleotide, a lipid, a
 CC saturated or unsaturated fatty acid, a diol, a carbohydrate, an aromatic
 CC compound, a vitamin, a cofactor, a polyketide, or an enzyme. The fine
 CC chemical production can be modulated. The presence of (I) or the SRT
 CC proteins (III) encoded by them are used for diagnosing the presence or
 CC activity of *Corynebacterium diptheriae*. (I), (II), (III) and host cells
 CC containing them can be used to map the genomes of organisms related to C.
 CC glutamicum, to identify and localise C. glutamicum sequences of interest,
 CC in evolutionary studies, in determination of SRT protein regions required
 CC for function, in modulating the SRT protein activity, and in modulating
 CC the activity of an SRT pathway. (II) are used to permit C. glutamicum to
 CC survive in an environment that is normally environmentally or chemically
 CC hazardous to it. (I) and protein molecules encoded by it increase the
 CC survival of C. glutamicum to chemical and environmental hazards and
 CC provide a means for continued growth and multiplication in large scale
 CC fermentative growth conditions. By increasing the growth rate or
 CC maintaining a normal growth rate in poor or toxic conditions, the yield,
 CC production and/or efficiency or production of fine chemicals from a
 CC culture may be increased

XX Sequence 67 AA:

Query Match 78.6%; Score 282; DB 4; Length 67;
 Best Local Similarity 77.6%; Pred. No. 7.5e-27;
 Matches 52; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

QY 1 MPOGTWVFAKGFPGFIAPEDGSADVFTYITIGTGFPTLBNOKVFEIGHSPGPQ 60
 1 MAOGTVMFNGKGFPGFIAPEDGSADVFTYITIGTGFPTLBNOKVFEIGHSPGPQ 60
 DB 61 ATGVRSI 67
 61 AAGVRL 67

RESULT 4
 ID AAB78911 standard; protein; 67 AA.

XX AAB78911;

XX 30-APR-2001 (first entry)

DB C. glutamicum SRT protein sequence SEQ ID NO:82.

XX *Corynebacterium glutamicum*; stress; resistance; tolerance; SRT;
 KW fine chemical production; organic acid; proteinogenic amino acid;
 KW nonproteinogenic amino acid; purine base; pyrimidine base; nucleoside;
 KW nucleotide; lipid; saturated fatty acid; unsaturated fatty acid; diol;
 KW carbohydrate; aromatic compound; vitamin; cofactor; polyketide; enzyme;
 KW evolutionary study; environmental hazard; fermentation.

XX *Corynebacterium glutamicum*.

XX MO200100804-A2.

XX 04-JAN-2001.

XX 23-JUN-2000; 2000MO-IB000922.

XX 25-JUN-1999; 99US-0141031P.

XX 01-JUL-1999; 99DE-01030429.

XX 01-JUL-1999; 99US-0142692P.

XX 08-JUL-1999; 99DE-01031413.

XX 08-JUL-1999; 99DE-01031457.

XX 09-JUL-1999; 99DE-01032209.

PR 09-JUL-1999; 99DE-01032230.
 PR 14-JUL-1999; 99DE-01032914.
 PR 27-AUG-1999; 99DE-01040764.
 PR 27-AUG-1999; 99US-0151214P.
 PR 31-AUG-1999; 99DE-01041382.
 PA (BAD1) BASF AG.
 XX
 XX Pompejus M, Kroegeer B, Schroeder H, Zelder O, Habernauer G;
 PI Lee H, Kim H;
 DR N-PSDB; AAF71024.
 XX WPI; 2001-061972/07.
 XX New isolated *Corynebacterium glutamicum* nucleic acid encoding a stress,
 PT tolerance or resistance protein, for production or modulation of
 PT production of fine chemicals, such as, e.g. amino acids, lipids,
 PT carbohydrates, or enzymes.

XX Claim 20; Page 230-231; 526pp; English.

CC AAF70984 to AAF71133 encode the *Corynebacterium glutamicum* stress,
 CC resistance, and tolerance (SRT) proteins given in AAB78871 to AAB79020.
 CC The C. glutamicum SRT genes (I) can be used in vectors (II) for
 CC expression in host cells and production of fine chemicals, such as, an
 CC organic acid, a proteinogenic or nonproteinogenic amino acid (preferred),
 CC a purine or pyrimidine base, a nucleoside, a nucleotide, a lipid, a
 CC saturated or unsaturated fatty acid, a diol, a carbohydrate, an aromatic
 CC compound, a vitamin, a cofactor, a polyketide, or an enzyme. The fine
 CC chemical production can be modulated. The presence of (I) or the SRT
 CC proteins (III) encoded by them are used for diagnosing the presence or
 CC activity of *Corynebacterium diptheriae*. (I), (II), (III) and host cells
 CC containing them can be used to map the genomes of organisms related to C.
 CC glutamicum, to identify and localise C. glutamicum sequences of interest,
 CC in evolutionary studies, in determination of SRT protein regions required
 CC for function, in modulating the SRT protein activity, and in modulating
 CC the activity of an SRT pathway. (II) are used to permit C. glutamicum to
 CC survive in an environment that is normally environmentally or chemically
 CC hazardous to it. (I) and protein molecules encoded by it increase the
 CC survival of C. glutamicum to chemical and environmental hazards and
 CC provide a means for continued growth and multiplication in large scale
 CC fermentative growth conditions. By increasing the growth rate or
 CC maintaining a normal growth rate in poor or toxic conditions, the yield,
 CC production and/or efficiency or production of fine chemicals from a
 CC culture may be increased

XX Sequence 67 AA:

Query Match 78.6%; Score 282; DB 4; Length 67;
 Best Local Similarity 77.6%; Pred. No. 7.5e-27;
 Matches 52; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

QY 1 MPOGTWVFAKGFPGFIAPEDGSADVFTYITIGTGFPTLBNOKVFEIGHSPGPQ 60
 1 MAOGTVMFNGKGFPGFIAPEDGSADVFTYITIGTGFPTLBNOKVFEIGHSPGPQ 60
 DB 61 ATGVRSI 67
 61 AAGVRL 67

RESULT 5

ID AAG89941 standard; protein; 67 AA.

XX AAG89941;

XX 26-SEP-2001 (first entry)

DB C. glutamicum protein fragment SEQ ID NO: 3695.

XX *Corynebacterium*; amino acid synthesis; vitamin; saccharide;
 KW organic acid synthesis.

OS Corynebacterium glutamicum.
 XX
 PN WO200100804-A2.
 XX
 PD 04-JAN-2001.
 XX
 PF 23-JUN-2000; 2000WO-IB000922.
 XX
 PR 25-JUN-1999; 99US-0141031P.
 PR 01-JUL-1999; 99DE-0103042P.
 PR 01-JUL-1999; 99US-0142692P.
 PR 08-JUL-1999; 99DE-01031413.
 PR 08-JUL-1999; 99DE-01031457.
 PR 08-JUL-1999; 99DE-01031541.
 PR 09-JUL-1999; 99DE-01032209.
 PR 09-JUL-1999; 99DE-01032230.
 PR 14-JUL-1999; 99DE-01032914.
 PR 27-AUG-1999; 99DE-01040764.
 PR 27-AUG-1999; 99US-0151214P.
 PR 31-AUG-1999; 99DE-01041382.
 XX
 PA (BADI) BASF AG.
 XX
 PI Pompejus M, Kroegeer B, Schroeder H, Zelder O, Habernauer G;
 PI Lee H, Kim H;
 DR N-PSDB; AAF71027.
 DR WPI; 2001-061972/07.
 PT New isolated Corynebacterium glutamicum nucleic acid encoding a stress,
 PT tolerance or resistance protein, for production or modulation of
 PT production of fine chemicals, such as, e.g. amino acids, lipids,
 PT carbohydrates, or enzymes.
 XX
 PS Claim 20; Page 234; 526pp; English.
 XX
 CC AAF70984 to AAF71133 encode the Corynebacterium glutamicum stress,
 CC resistance, and tolerance (SRT) proteins given in AAB78871 to AAB79020.
 CC The C. glutamicum SRT genes (I) can be used in vectors (II) for
 CC expression in host cells and production of fine chemicals, such as, an
 CC organic acid, a proteinogenic or nonproteinogenic amino acid (preferred),
 CC a putine or pyrimidine base, a nucleoside, a nucleotide, a lipid, a
 CC saturated or unsaturated fatty acid, a diol, a carbohydrate, an aromatic
 CC compound, a vitamin, a cofactor, a polyketide, or an enzyme. The fine
 CC chemical production can be modulated. The presence of (I) or the SRT
 CC proteins (III) encoded by them are used for diagnosing the presence or
 CC activity of Corynebacterium diptheriae. (I), (II), (III) and host cells
 CC containing them can be used to map the genomes of organisms related to C.
 CC glutamicum, to identify and localise C. glutamicum sequences of interest,
 CC in evolutionary studies, in determination of SRT protein regions required
 CC for function, in modulating the SRT protein activity, and in modulating
 CC the activity of an SRT pathway. (II) are used to permit C. glutamicum to
 CC survive in an environment that is normally environmentally or chemically
 CC hazardous to it. (I) and protein molecules encoded by it increase the
 CC survival of C. glutamicum to chemical and environmental hazards and
 CC provide a means for continued growth and multiplication in large scale
 CC fermentative growth conditions. By increasing the growth rate or
 CC maintaining a normal growth rate in poor or toxic conditions, the yield,
 CC production and/or efficiency or production of fine chemicals from a
 CC culture may be increased
 XX
 SQ Sequence 67 AA;
 Query Match 76.6%; Score 275; DB 4; Length 67;
 Best Local Similarity 77.6%; Pred. No. 5.5e-26;
 Matches 52; Conservative 4; Mismatches 11; Indels 0; Gaps 0;
 QY 1 MPOGTVMFNAKGFPIAPEDGSADVFNHYTEIOGTGPTLLENOKVEFEIGHSPKGPQ 60
 DB 1 MPOGTVMFNAKGFPIAPEDGSADVFNHYTEIOGTGPTLLENOKVEFEIGHSPKGPQ 60
 QY 61 ATGVRSI 67
 DB 61 ATGVRSI 67

DB 61 AQAIVRAI 67
 RESULT 8
 AAG90091
 ID AAG90091 standard; protein; 67 AA.
 XX
 AC AAG90091;
 XX
 DT 26-SEP-2001 (first entry)
 XX
 DB C glutamicum protein fragment SEQ ID NO: 3845.
 XX
 KM Coryneform bacterium; amino acid synthesis; vitamin; saccharide;
 KM organic acid synthesis.
 XX
 OS Corynebacterium glutamicum.
 XX
 PN BP108790-A2.
 XX
 PD 20-JUN-2001.
 XX
 PF 18-DEC-2000; 2000EP-00127688.
 XX
 PR 16-DEC-1999; 99JP-00377484.
 PR 07-APR-2000; 2000JP-00159162.
 PR 03-AUG-2000; 2000JP-00280988.
 XX
 PA (KYOW) KYOWA HAKKO KOGYO KK.
 XX
 PI Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
 PI Tateishi N, Senoh A, Ikeda M, Ozaki A;
 DR N-PSDB; AAB65310.
 DR WPI; 2001-376931/40.
 XX
 PT Novel polynucleotides derived from Coryneform bacteria, for identifying
 PT mutation point of a gene, measuring expression of a gene, analyzing
 PT expression profile or pattern of a gene and identifying homologous gene.
 XX
 PS Claim 17; SEQ ID NO 3845; 246pp + Sequence listing; English.
 XX
 CC The present invention provides a number of nucleotide and protein
 CC sequences from the Coryneform bacterium Corynebacterium glutamicum. These
 CC are useful for identifying the mutation point of a gene derived from a
 CC mutant of coryneform bacterium, measuring expression amount and analysing
 CC the expression profile or expression pattern of a gene derived from
 CC Coryneform bacterium, and identifying a homologue of a gene derived from
 CC coryneform bacterium. Coryneform bacteria are useful for producing amino
 CC acids, nucleic acids, vitamins, saccharides and organic acids,
 CC particularly L-lysine. The present sequence is a protein described in the
 CC exemplification of the invention. Note: The sequence data for this patent
 CC did not form part of the printed specification, but was obtained in
 CC electronic format directly from the European Patent Office
 XX
 SQ Sequence 67 AA;
 Query Match 76.6%; Score 275; DB 4; Length 67;
 Best Local Similarity 77.6%; Pred. No. 5.5e-26;
 Matches 52; Conservative 4; Mismatches 11; Indels 0; Gaps 0;
 QY 1 MPOGTVMFNAKGFPIAPEDGSADVFNHYTEIOGTGPTLLENOKVEFEIGHSPKGPQ 60
 DB 1 MPOGTVMFNAKGFPIAPEDGSADVFNHYTEIOGTGPTLLENOKVEFEIGHSPKGPQ 60
 QY 61 AQAIVRAI 67
 DB 61 AQAIVRAI 67
 RESULT 9
 AB034189
 ID AB034189 standard; protein; 67 AA.

XX	ABU34189;
AC	
XX	19-JUN-2003 (first entry)
DT	
DE	Protein encoded by Prokaryotic essential gene #19716.
XX	
KM	Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX	
OS	Mycobacterium avium.
XX	
PN	WO2002717183-A2.
PD	
XX	03-OCT-2002.
PF	
XX	21-MAR-2002; 2002MO-US009107.
PR	
XX	21-MAR-2001; 2001US-00815242.
PR	06-SEP-2001; 2001US-00948993.
PR	25-OCT-2001; 2001US-0342923P.
PR	08-FEB-2002; 2002US-00072851.
PR	06-MAR-2002; 2002US-0362699P.
XX	
PA	(ELITRA) ELITRA PHARM INC.
P1	
F1	Wang L, Zamudio C, Malone C, Hasebeck R, Ohlsen KU, Zykling JW;
DR	Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX	
DR	WPI; 2003-029926/02.
XX	N-Psdb; ACA38059.
PT	
PT	New antisense nucleic acids, useful for identifying proteins or screening
PT	for homologous nucleic acids required for cellular proliferation to
PT	isolate candidate molecules for rational drug discovery programs.
PS	
XX	Claim 25; SEQ ID NO 62113; 1766bp; English.
CC	The invention relates to an isolated nucleic acid comprising any one of
CC	the 6213 antisense sequences given in the specification where expression
CC	of the nucleic acid inhibits proliferation of a cell. Also included are:
CC	(1) a vector comprising a promoter operably linked to the nucleic acid
CC	encoding a polypeptide whose expression is inhibited by the antisense
CC	nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC	polypeptide or its fragment whose expression is inhibited by the
CC	antisense nucleic acid; (4) an antibody capable of specifically binding
CC	the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC	proliferation or the activity of a gene in an operon required for
CC	proliferation; (7) identifying a compound that influences the activity of
CC	the gene product or that has an activity against a biological pathway
CC	required for proliferation, or that inhibits cellular proliferation; (8)
CC	identifying a gene required for cellular proliferation or the biological
CC	pathway in which a proliferation-required gene or its gene product lies
CC	or a gene on which the test compound that inhibits proliferation of an
CC	organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC	compound's activity; (11) a culture comprising strains in which the gene
CC	product is overexpressed or underexpressed; (12) determining the extent
CC	to which each of the strains is present in a culture or collection of
CC	strains; or (13) identifying the target of a compound that inhibits the
CC	proliferation of an organism. The antisense nucleic acids are useful for
CC	identifying proteins or screening for homologous nucleic acids required
CC	for cellular proliferation to isolate candidate molecules for rational
CC	drug discovery programs, or for screening homologous nucleic acids
CC	required for proliferation in cells other than S. aureus, S. typhimurium,
CC	K. pneumoniae or P. aeruginosa. The present sequence is encoded by one of
CC	the target prokaryotic essential genes. Note: The sequence data for this
CC	patent did not form part of the printed specification, but was obtained
CC	in electronic format directly from WIPO at
CC	ftp.wipo.int/pub/published_pct_sequences
XX	
SQ	Sequence 67 AA:
Query Match	71.3%; Score 256; DB 6; Length 67;
Best Local Similarity	64.2%; Pred. No. 1,2e-23;

Matches	43;	Conservative	14;	Mismatches	10;	Indels	0;	Gaps	0;
Qy		1	MPOGTVMFNAEKGFEGFIAPBDGSADVFWHTHTIGTGFFLTLENQVFEIGHSPKGPQ						60
Db		1	MAOQTVMFNGEKGFEGFITPDDGTDFVHYSIQSGVSLDENQVQPDVEGAKGPQ						60
Qy		61	ATGVRS						67
Db		61	AVGVSTV						67
RESULT 10									
		ID	ADH86286						standard; protein: 88 AA.
		ADH86286;							
		22-APR-2004							(first entry)
		Enterococcus faecalis polypeptide #766.							
		Enterococcus faecalis infection; transcription regulatory element;							
		antibacterial.							
		Enterococcus faecalis.							
		US6617156-B1.							
		09-SEP-2003.							
		13-AUG-1998;							98US-00134000.
		15-AUG-1997;							97US-0055778P.
		(DOUC/) DOUCETTE-STAMM L A.							
		(BUSH/) BUSH D.							
		Doucette-Stamm LA, Bush D;							
		WP1; 2003-895394/82.							
		N-PSDB; ADH82881.							
		New nucleic acid comprising a sequence encoding an Enterococcus faecalis							
		polypeptide, useful for preparing a composition for diagnosing or							
		treating E. faecalis infection.							
		Disclosure; SEQ ID NO 4171; 193bp; English.							
		The invention relates to Enterococcus faecalis polynucleotides and							
		polypeptides. The invention also relates to a recombinant expression							
		vector comprising a polynucleotide operably linked to a transcription							
		regulatory element, a cell comprising a recombinant vector, a method for							
		producing an E. faecalis polypeptide, an isolated nucleic acid comprising							
		a sequence not given in the specification, a recombinant vector							
		comprising the nucleic acid and a cell comprising the recombinant vector.							
		The polynucleotides can be used to detect the presence of E. faecalis in							
		a sample. The sequences are useful for preparing a composition for							
		diagnosing or treating Enterococcus faecalis infection. This sequence							
		represents an E. faecalis polypeptide of the invention.							
		Sequence 88 AA;							
		Query Match							
		Best Local Similarity							
		Matches							
		44;							
		Conservative							
		7;							
		Mismatches							
		12;							
		Indels							
		1;							
		Gaps							
		1;							
		Score							
		63.44;							
		Score							
		227.5;							
		DB							
		7;							
		Length							
		88;							
		Query Match							
		Best Local Similarity							
		Matches							
		44;							
		Conservative							
		7;							
		Mismatches							
		12;							
		Indels							
		1;							
		Gaps							
		1;							
		Score							
		63.44;							
		Score							
		227.5;							
		DB							

RESULT 11

ID AEM78962 standard; protein: 70 AA.

XX AEM78962;

XX 15-JAN-2004 (first entry)

XX Escherichia coli cold shock protein CspG.

XX Cold shock; CspG.

XX Escherichia coli.

XX W02003074657-A2.

XX 12-SEP-2003.

XX 25-FEB-2003; 2003MO-US005531.

XX 01-MAR-2002; 2002US-0361069P.

XX 14-AUG-2002; 2002US-0402921P.

XX (TAKA-) TAKARA BIO INC.

XX (MCIN/) MCINNIS P.

XX Inouye M, Phadtare S, Xia B, Qing G, Ke H;

XX WPI; 2003-731677/69.

XX N-PSDB; ACF79942.

XX New DNA encoding heterologous protein, useful particularly for producing

XX isotopically labeled proteins, includes a cold-shock inducible promoter.

XX Disclosure; Page 73; 83pp; English.

XX The present sequence is the protein sequence of Escherichia coli cold shock protein CspG. The invention provides a DNA molecule or vector and a host cell containing the DNA molecule or vector for use in the production of a heterologous polypeptide under conditions that elicit a cold shock response in the host cells. The DNA molecule and vector include a nucleotide sequence encoding the heterologous polypeptide and a promoter and 5' untranslated region (UTR) from a cold shock inducible gene, such as cspG, which directs its expression. An AT-rich sequence that enhances translation under cold shock inducible conditions is present either in the coding sequence of the heterologous polypeptide or in an additional element inserted between the coding sequence and the cold shock inducible promoter and 5'-UTR. The DNA molecule preferably also includes the cspA 3'-UTR sequence. Inducing expression by cold shock eliminates the need for chemical induction, and under cold conditions all ribosomes in the host are forced to express the single target protein, which simplifies purification. When used to prepare isotopically labelled proteins, the culture size, and thus amount of expensive labelled reagent, can be reduced. The method can also be used to produce proteins native to cold-climate organisms at high yield

XX Sequence 70 AA;

SQ Query Match 62.1%; Score 223; DB 7; Length 70;

Best Local Similarity 65.6%; Pred. No. 1.5e-19;

Matches 42; Conservative 5; Mismatches 17; Indels 0; Gaps 0;

QY 4 GTWKMPNAKGFSPFIAPEDGSADVFAHYTETIGTGFTLEBNQKVEFEIGHSPKGPATG 63

DB 7 GLVKMFADKRGFPITPDGSKOVFVHFTAIQSNBFTLNMNQKVEFSIEGCGRGPAAN 66

QY 64 VRSL 67

DB 67 VVTL 70

RESULT 12

ID ADZ21446 standard; protein: 70 AA.

XX ADZ21446;

XX 30-JUN-2005 (first entry)

XX Escherichia coli cold shock protein.

XX Abiotic stress tolerance; transgenic plant; crop improvement;

XX heat tolerance; cold tolerance; salt tolerance; drought tolerance;

XX cold shock protein.

XX Escherichia coli.

XX W02005033318-A2.

XX 14-APR-2005.

XX 29-SEP-2004; 2004MO-US031856.

XX 29-SEP-2003; 2003US-0506717P.

XX 17-DEC-2003; 2003US-0530453P.

XX (MONS) MONSANTO TECHNOLOGY LLC.

XX Fernandez M;

XX WPI; 2005-285434/29.

XX N-PSDB; ADZ21445.

XX New recombinant DNA molecule comprising a DNA polynucleotide comprising a

XX promoter and a DNA that encodes a cold shock protein, useful for

XX producing plants with increased tolerance to abiotic stress.

XX Claim 4; SEQ ID NO 19; 169pp; English.

XX The present sequence is the protein sequence of a cold shock protein (Csp) from Escherichia coli. The invention provides a plant that shows improved abiotic stress tolerance as a result of the expression of a Csp in its cells. In one embodiment, a polynucleotide encoding a Csp is operably linked to a plant promoter and plant terminator. The polynucleotide preferably encodes a protein comprising a specified Csp conserved motif ADZ21430, particularly Bacillus subtilis CspB ADZ21429 or its homologues, Escherichia coli CspA ADZ21428 or its homologues, a Csp from Agrobacterium tumefaciens or other Csp protein of the invention, such as the present Csp. The promoter is selected from an inducible, constitutive, temporally-regulated, developmentally-regulated, tissue-specific, drought inducible, water deficit inducible or tissue-specific promoter. An abiotic stress-tolerant transgenic plant that has been transformed with a DNA molecule that expresses a Csp is claimed. The plant may be a crop plant, a monocot or dicot, and is especially selected from soybean, corn, canola, rice, cotton, barley, oat, turf grass, cotton and wheat. The transgenic plant shows improved heat tolerance, cold tolerance, salt tolerance, drought tolerance and survival after cold shock than the corresponding non-transformed plant.

XX Sequence 70 AA;

SQ Query Match 62.1%; Score 223; DB 9; Length 70;

Best Local Similarity 65.6%; Pred. No. 1.5e-19;

Matches 42; Conservative 5; Mismatches 17; Indels 0; Gaps 0;

QY 4 GTWKMPNAKGFSPFIAPEDGSADVFAHYTETIGTGFTLEBNQKVEFEIGHSPKGPATG 63

DB 7 GLVKMFADKRGFPITPDGSKOVFVHFTAIQSNBFTLNMNQKVEFSIEGCGRGPAAN 66

QY 64 VRSL 67

DB 67 VVTL 70

RESULT 13

ABO66959 standard; protein; 131 AA.

ABO66959;

29-JUL-2004 (first entry)

Klebsiella pneumoniae polypeptide seqid 13476.

Recombinant expression vector; transcription regulatory element;
Klebsiella pneumoniae protein; antibacterial; Vaccine.

Klebsiella pneumoniae.

US610836-B1.

26-AUG-2003.

27-JAN-2000; 2000US-00489039.

29-JUN-1999; 99US-0117747P.

(GENO-) GENOME THERAPEUTICS CORP.

Breton GL, Osborne M;

WPI; 2003-895346/82.

N-PSDB; ABD00530.

New nucleic acid encoding a *Klebsiella pneumoniae* polypeptide, useful for preparing a vaccine composition against *Klebsiella pneumoniae*.

Disclosure; SEQ ID NO 13476; 932bp; English.

The invention describes a new isolated nucleic acid encoding a *Klebsiella pneumoniae* polypeptide. Also described are: a recombinant expression vector comprising the nucleic acid, operably linked to a transcription regulatory element; and a cell comprising the recombinant expression vector. The nucleic acid is useful for preparing a vaccine composition against *Klebsiella pneumoniae*. This is the amino acid sequence of a *Klebsiella pneumoniae* polypeptide of the invention

Sequence 131 AA;

Query Match 61.8%; Score 222; DB 7; Length 131;

Best Local Similarity 63.1%; Pred. No. 4,4e-19; Indels 0; Gaps 0;

Matches 41; Conservative 7; Mismatches 17; Indels 0; Gaps 0;

3 OGTVKFNNAKGFPGFTAPEDGSADVFNHYTEIOGTGRTLENNQKVFETGHSKPGQAT 62

67 KGNVKNFNSKGFPTPEBGSKDVFVHFSAIISNGKTLAEGORVFEITTNAGKPSAA 126

63 GVRSI 67

127 NVIAL 131

RESULT 14

ADH54477

ADH54477 standard; protein; 70 AA.

ADH54477;

25-MAR-2004 (first entry)

Escherichia coli cspg protein.

Cold shock inducible gene; physiological stress; cs; cspg.

Escherichia coli.

Escherichia coli.

Escherichia coli.

US610533-B1.

26-AUG-2003.

01-MAR-2000; 2000US-00516667.

01-MAR-2000; 2000US-00516667.

(UYNE-) UNIV NEW JERSEY MEDICINE & DENTISTRY.

Inouye M, Wang N, Yamanaka K;

WPI; 2003-895341/82.

New nucleic acid comprising the 5'-UTR of cspl operably linked to a non-native promoter, useful for regulating expression of cold shock inducible genes to elicit the cold shock response in bacterium under conditions of physiological stress.

Example 3; SEQ ID NO 80; 86bp; English.

The invention relates to a nucleic acid molecule which prolongs the expression of cold shock inducible genes under conditions that elicit the cold shock (cs) response in a bacterium, comprises the 5'-UTR of cspl operably linked to a non-native promoter. The nucleic acid is useful for regulating expression of cold shock inducible genes to elicit the cold shock response in bacterium under conditions of physiological stress. The present sequence is *Escherichia coli* cspg protein used in the exemplification of the invention.

Sequence 70 AA;

Query Match 61.6%; Score 221; DB 7; Length 70;

Best Local Similarity 64.1%; Pred. No. 2,6e-19; Indels 0; Gaps 0;

Matches 41; Conservative 6; Mismatches 17; Indels 0; Gaps 0;

4 GTYKFNNAKGFPGFTAPEDGSADVFNHYTEIOGTGRTLENNQKVFETGHSKPGQAT 63

7 GLVKNFNAKGFPGFTAPEDGSADVFNHYTEIOGTGRTLENNQKVFETGHSKPGQAT 66

64 VRSI 67

67 VVTL 70

RESULT 15

ABB07338

ABB07338 standard; protein; 66 AA.

ABB07338;

09-APR-2002 (first entry)

S. clavuligerus CspA-like protein SC7.0.

Cold-shock induced protein; Csp; prokaryotic; stress; desaturase; SC7.0.

Streptomyces clavuligerus.

US6333191-B1.

25-DEC-2001.

03-FEB-1998; 98US-00017754.

13-FEB-1989; 89US-00310332.

09-MAR-1992; 92US-00852013.

01-MAR-1994; 94US-00203806.

(UYNE-) UNIV NEW JERSEY MEDICINE & DENTISTRY.

Inouye M, Jones P, Etchegaray J, Uiang W, Pollitt NS;

Goldstein J;

Goldstein J;

Goldstein J;

XX WPI; 2002-129568/17.
DR

PT DNA molecules useful for encoding proteins under stress or cold-induced
PT conditions, comprising a prokaryotic promoter that is induced under
PT conditions eliciting a cold shock response.

PS Disclosure; Fig 10; 33pp; English.

The invention relates to a family of stimuli-induced in particular stress or cold-shock induced genes and proteins. An isolated DNA molecule (I) is provided that comprises a prokaryotic promoter that is induced under conditions eliciting a cold shock response (Csp), or the promoter that is repressed under conditions eliciting Csp, where the conditions comprise a decrease in temperature below the normal physiological growth temperature of the prokaryotic cell, preferably bacterium. (I) is useful for encoding proteins under stimuli-induced, particularly stress or cold-shock induced conditions. The encoded proteins are useful in staining DNA and in the stabilization of DNA and RNA, and thus increasing the efficiency of DNA and RNA in various *in vitro* reactions. The encoded protein is also useful as deaturase to denature DNA irrespective of temperature. The proteins are also useful for commercial purposes. The present sequence represents a *S. clavuligerus* CspA-like protein SC7.0

SQ Sequence 66 AA;

Query Match	60.9%	Score 218.5	DB 5	Length 66
Best Local Similarity	65.6%	Pred. No. 4.9e-19		
Matches 42; Conservative	6	Mismatches 15	Indels 1	Gaps 1

Qy 1 MPQGYKMFNAEKGFPIAPEDDSADVPHYTYELQGTGRITLEENQKVBEIIGHSGKGPQ 60
| | | | | : | | | | | : | | : | |
Db 1 MATGYKMFNAEKGFPIAQDGGSDPVFHYSAINATGRSLBEQVNVNDVTHG-EGPQ 59

1 MATGTVKWFNAEKGFIAQDGGP DVFVHYSAINATGFRSLBENQVNVFDVTHG-EGPQ 59

QY 61 ATGV 64

Db 60 252V 63

Search completed: April 14, 2006, 17:18:45
Job time : 28.9599 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 14, 2006, 17:19:08 ; Search time 3.92873 Seconds

(without alignments)
1640.866 Million cell updates/sec

Title: US-10-620-246-147

Perfect score: 359

Sequence: 1 MFGQTVKMFNAEKGFPIADP.....VEPFIHSPKPGQATVRSLL 67

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	359	100.0	67	2 A70564	probable csppA prot
2	352	98.1	67	2 F86933	cold shock-like pr
3	223	62.1	70	2 A90772	cold shock-like pr
4	223	62.1	70	2 D64840	cold shock-like pr
5	222	61.8	69	2 A84987	cold shock-like pr
6	222	61.8	69	2 A80579	cold shock-like pr
7	221	61.6	69	2 F84967	cold shock-like pr
8	221	61.6	70	2 B85634	homolog of Salmo
9	220	61.3	69	2 B85632	cold shock protein
10	220	61.3	69	2 S49050	cold shock protein
11	220	61.3	69	2 F90711	cold shock protein
12	219	61.0	70	2 AD0755	cold shock protein
13	218.5	60.9	66	2 S26378	cold shock protein
14	217	60.4	69	2 A10212	cold shock protein
15	217	60.4	70	2 C64910	cold shock protein
16	216	60.2	69	2 AB0727	cold shock-like pr
17	216	60.2	69	2 B90945	cold shock protein
18	216	60.2	69	2 S43618	cold shock protein
19	216	60.2	69	2 H85793	cold shock protein
20	215	59.9	69	2 AG0316	probable cold shoc
21	211	58.8	70	1 OCBCU	major cold shock p
22	211	58.8	70	2 AG0981	cold shock protein
23	211	58.8	70	2 A91184	cold shock protein
24	211	58.8	70	2 B86030	cold shock protein
25	210	58.5	70	2 D70412	cold shock protein
26	209.5	58.4	65	2 H86645	cold shock protein
27	207.5	57.8	66	2 AD1245	cold shock protein
28	207.5	57.8	66	2 AH1607	cold shock protein
29	207	57.7	67	2 G82568	cold shock protein

30	206.5	57.5	66	2 AH1326	major cold-shock p
31	206.5	57.5	66	2 AB1698	major cold-shock p
32	206.5	57.5	66	2 I40158	major cold shock p
33	206	57.4	70	2 AB0202	cold shock protein
34	204	56.8	70	2 AC0324	cold shock protein
35	202.5	56.4	65	2 B84101	cold-shock protein
36	202.5	56.4	66	2 I40380	major cold shock p
37	201	56.0	67	2 T34593	cold-shock domain
38	201	56.0	70	2 AB0443	major cold shock p
39	200.5	55.8	66	2 B69608	cold shock protein
40	200.5	55.8	69	2 A82489	cold shock DNA-bin
41	200	55.7	70	2 AD0443	major cold shock p
42	199.5	55.6	66	2 A89853	cold-shock protein
43	199.5	55.6	66	2 B89917	major cold shock p
44	199.5	55.6	67	2 A45723	major cold shock p
45	199	55.4	67	2 T42055	cold shock protein

ALIGNMENTS

RESULT 1
A70564
probable csppA protein - Mycobacterium tuberculosis (strain H37RV)
C/Species: Mycobacterium tuberculosis
C/Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 31-Dec-2004
C/Accession: A70564
R/Col: S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, R.; Andreu, M.A.; Rogers, J.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Connor, R.; Davies, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, R.; Andreu, M.A.; Rogers, J.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Nature 393, 537-544, 1998
A/Authors: Sgares, R.; Sultam, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A/Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A/Reference number: A70500; PMID:9825987; PMID:9634230
A/Accession: A70564
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-67 <CD>
A/Cross-references: UNIPROT:O06360; UNIPARC:UP1000012856D; GB:Z95436; GB:AL123456; NID: C/Experimental source: strain H37RV
C/Genetics:
A/Gene: csppA
C/Superfamily: cold shock protein, csppA type; cold shock domain homology
F/4-64/Domain: cold shock domain homology <CSD>
Query Match 100.0%; Score 359; DB 2; Length 67;
Best Local Similarity 100.0%; Pred. No. 3.7e-35;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Cy 1 MFGQTVKMFNAEKGFPIADPDSADVFTYTRIGTGFRTLBENQVBEFGHSPKGPQ 60
Db 1 MFGQTVKMFNAEKGFPIADPDSADVFTYTRIGTGFRTLBENQVBEFGHSPKGPQ 60
Cy 61 ATGVRSL 67
Db 61 ATGVRSL 67
RESULT 2
F86933
probable cold shock protein [imported] - Mycobacterium lprae
C/Species: Mycobacterium lprae
C/Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 31-Dec-2004
C/Accession: F86933
R/Col: S.T.; Bismeyer, K.; Parkhill, J.; James, K.D.; Thomson, N.R.; Wheeler, P.R.; H. R.; Davies, R.M.; Devlin, K.; Duthey, S.; Feltwell, T.; Frazer, A.; Hamlin, N.; Holroyd, S.; Nature 409, 1007-1011, 2001
A/Authors: Rutter, S.; Seeger, K.; Simon, S.; Skelton, M.; Skelton, J.; Squares, R.; S. A/Title: Massive gene decay in the leprosy bacillus.
A/Reference number: A86509; PMID:21128732; PMID:11234002
A/Accession: F86933
A/Status: preliminary

A: CIOB8-References: UNIPARC:DP10000035A1/6; GB:AMJ33582; EBLN:CA003106.1
C: Geneticks:
A: Gene: STY0678
C: superfamily: cold shock protein, Cspa type; cold shock domain homology

Query Match 61.8%; Score 222; DB 2; Length 69;
 Best Local Similarity 63.1%; Pred. No. 4.3e-19;
 Matches 41; Conservative 7; Mismatches 17; Indels 0; Gaps 0;

Db 3 OGTVKMFNAKGGFGFIAPEDSGADVFAHYTEIQGTGFRITLBNQKVFEEIGHSPKGPQAT 62
 5 KGVNKMNFESKGGFIFIPEDSGKDVFAHFSALQTNKGKTLAEGQKVFEEITNGAKGFSNA 64

QY 63 GVRSL 67
 65 NVNVL 69

Db 63 GVRSL 67
 65 NVNVL 69

RESULT 7
 P84967
 cold shock-like protein cspp [imported] - Buchnera sp. (strain APS)
 C:Species: Buchnera sp.
 C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Dec-2004
 C/Accession: F84967
 R/Shigenobu, S.; Watanabe, H.; Hattori, M.; Sakaki, Y.; Ishikawa, H.
 Nature 407, 81-86, 2000
 A>Title: Genome sequence of the endocellular bacterial symbiont of aphids Buchnera sp. A
 A/Reference number: A84930; MUID:20445173; PMID:10993077
 A/Accession: F84967
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-69 <STO>
 A/Cross-references: UNIPARC:UPI0000055539; GB:AP000398; GSPDB:GM00144
 C/Experimental source: strain APS
 C/Genetics:
 A/Gene: csppC; BU322
 C/Superfamily: cold shock protein, Cspp type; cold shock domain homology

Query Match 61.6%; Score 221; DB 2; Length 69;
 Best Local Similarity 64.6%; Pred. No. 5.6e-19;
 Matches 42; Conservative 4; Mismatches 19; Indels 0; Gaps 0;

QY 3 OGTVKMFNAKGGFGFIAPEDSGADVFAHYTEIQGTGFRITLBNQKVFEEIGHSPKGPQAT 62
 5 KGVNKMNFESKGGFIFIPEDSGKDVFAHFSALQTNKGKTLAEGQKVFEEITNGAKGFSNA 64

Db 5 KGVNKMNFESKGGFIFIPEDSGKDVFAHFSALQTNKGKTLAEGQKVFEEITNGAKGFSNA 64

QY 63 GVRSL 67
 65 NVNVL 69

Db 63 GVRSL 67
 65 NVNVL 69

RESULT 8
 B85634
 homolog of Salmonella cold shock protein [imported] - Escherichia coli (strain O157:H7,
 C:Species: Escherichia coli
 C>Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 31-Dec-2004
 C/Accession: B85634
 R/Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
 Iller, L.; Grobeck, B.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
 Nature 409, 529-533, 2001
 A>Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
 A/Reference number: A85480; MUID:21074935; PMID:11206551
 A/Accession: B85634
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-70 <STO>
 A/Cross-references: UNIPARC:UPI0000165753; GB:AE005174; NID:g12514257; PIDN:ANG55537.1;
 A/Experimental source: strain O157:H7, substrain EDL933
 C/Genetics:
 A/Gene: csppG
 C/Superfamily: cold shock protein, Cspp type; cold shock domain homology

Query Match 61.6%; Score 221; DB 2; Length 70;
 Best Local Similarity 67.2%; Pred. No. 5.7e-19;
 Matches 41; Conservative 4; Mismatches 16; Indels 0; Gaps 0;

QY 4 GTVKMFNAKGGFGFIAPEDSGADVFAHYTEIQGTGFRITLBNQKVFEEIGHSPKGPQAT 63
 5 KGVNKMNFESKGGFIFIPEDSGKDVFAHFSALQTNKGKTLAEGQKVFEEITNGAKGFSNA 64

Db 7 GLVKMFNADKGGFGFIAPEDSGKDVFAHFSALQTNKGKTLAEGQKVFEEIGHSPKGPQAT 66
 5 KGVNKMNFESKGGFIFIPEDSGKDVFAHFSALQTNKGKTLAEGQKVFEEITNGAKGFSNA 64

QY 64 V 64
 67 V 67

Db 67 V 67

RESULT 9
 B85662
 cold shock protein [imported] - Escherichia coli (strain O157:H7, substrain EDL933)
 C:Species: Escherichia coli
 C>Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 31-Dec-2004
 C/Accession: B85662
 R/Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
 Iller, L.; Grobeck, B.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
 Nature 409, 529-533, 2001
 A>Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
 A/Reference number: A85480; MUID:21074935; PMID:11206551
 A/Accession: B85662
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-69 <STO>
 A/Cross-references: UNIPARC:UPI00000023B; GB:AE005174; NID:g12513525;
 A/Experimental source: strain O157:H7, substrain EDL933
 C/Genetics:
 A/Gene: csppE
 C/Superfamily: cold shock protein, Cspp type; cold shock domain homology

Query Match 61.3%; Score 220; DB 2; Length 69;
 Best Local Similarity 63.1%; Pred. No. 7.4e-19;
 Matches 41; Conservative 7; Mismatches 17; Indels 0; Gaps 0;

QY 3 OGTVKMFNAKGGFGFIAPEDSGADVFAHYTEIQGTGFRITLBNQKVFEEIGHSPKGPQAT 62
 5 KGVNKMNFESKGGFIFIPEDSGKDVFAHFSALQTNKGKTLAEGQKVFEEITNGAKGFSNA 64

Db 5 KGVNKMNFESKGGFIFIPEDSGKDVFAHFSALQTNKGKTLAEGQKVFEEITNGAKGFSNA 64

QY 63 GVRSL 67
 65 NVNVL 69

Db 63 GVRSL 67
 65 NVNVL 69

RESULT 10
 S49050
 cold shock protein homolog cspp - Escherichia coli (strain K-12)
 C:Species: Escherichia coli
 C>Date: 29-Nov-1995 #sequence_revision 01-Mar-1996 #text_change 31-Dec-2004
 C/Accession: S49050; B64796
 R/Yamanaka, K.; Mitani, T.; Ogura, T.; Niki, H.; Hiraga, S.
 Mol. Microbiol. 13, 301-312, 1994
 A>Title: Cloning, sequencing, and characterization of multicopy suppressors of a mukB m
 A/Reference number: S49048; MUID:95075315; PMID:7984109
 A/Accession: S49050
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-69 <YNA>
 A/Cross-references: UNIPROT:P36997; UNIPARC:UPI00000023B; EMBL:D28497; NID:g460700; PI
 A/Experimental source: strain K-12
 R/Battner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; C
 Science 277, 1453-1462, 1997
 A>Title: The complete genome sequence of Escherichia coli K-12.
 A/Reference number: A64720; MUID:97426617; PMID:9278503
 A/Accession: B64796
 A/Status: nucleic acid sequence not shown; translation not shown
 A/Molecule type: DNA
 A/Residues: 1-69 <BLAT>
 A/Cross-references: UNIPARC:UPI00000023B; GB:AE000167; GB:U00096; NID:g1786836; PIDN:A
 A/Experimental source: strain K-12, substrain MG1655
 C/Genetics:
 A/Gene: csppE, mcmC, glcA
 C/Superfamily: cold shock protein, Cspp type; cold shock domain homology
 C/Keywords: DNA binding; transcription regulation
 F/6-66/Domain: cold shock domain homology <CSD>

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 14, 2006, 17:04:10 ; Search time 24.7208 Seconds
(without alignments)
1912.171 Million cell updates/sec

Title: US-10-620-246-147

Perfect score: 359
Sequence: 1 MPOGTYKMFNAEKGFRIAP.....VEPFIHSPKPGQATGVRL 67

Scoring table: BL0SUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : UniProt_05.80:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	359	100.0	67	1	CSPA_MYCBO
2	359	100.0	67	1	CSPA_MYCBO
3	359	100.0	67	2	0744B8_MYCPA
4	352	98.1	67	2	069550_MYCBP
5	341	95.0	67	2	09KGM0_MYCSM
6	319	88.9	68	2	0522Y3_NOCFA
7	302	84.1	67	2	06NUS3_CORDI
8	296	82.5	67	2	04JSS1_CORJK
9	289	80.5	251	2	08FLY0_CORRF
10	282	78.6	67	2	08NRY7_CORCL
11	275	76.3	67	2	08NTU9_CORCL
12	274	76.3	67	2	08FSS6_CORRF
13	266	74.1	67	2	04NFR7_9MICC
14	260	72.4	67	1	CSP_ARTGO
15	258	71.3	67	2	04NLE1_9MICC
16	256	71.3	67	2	0743B6_MYCPA
17	254	70.8	67	2	06AC21_LEIXX
18	249	69.4	67	2	06N4H9_9MICC
19	244.5	68.1	73	2	06AC11_LEIXX
20	242	67.4	67	2	05YU56_NOCFA
21	239	66.6	67	1	CSPA_MITLU
22	238	66.3	68	2	05YR77_NOCFA
23	235	65.5	67	2	082D26_STRAW
24	233.5	65.0	67	2	09EW90_9ACTO
25	233.5	65.0	67	2	09KXU6_STRCO
26	233.5	65.0	68	2	09EUN8_STRAY
27	233	64.9	68	2	07WXR6_BORBP
28	233	64.9	68	2	07WXR5_BORBP
29	232	64.6	67	2	09LOT9_STRCO
30	230.5	64.2	68	2	082G55_STRAW
31	230	64.1	65	2	06YANO_9ACTO

32	229	63.8	67	2	08XTV2_RALCO
33	227.5	63.4	66	2	0822X2_ENTFA
34	226.5	63.1	66	2	08R385_THENTN
35	225	62.7	66	2	07VXJ0_BORBP
36	224	62.4	67	1	CSPA_BORBR
37	224	62.4	67	1	CSPA_BORPA
38	224	62.4	67	1	CSPA_BORPE
39	224	62.4	68	2	04USH0_XANCP
40	224	62.4	68	2	07CLT8_XANCP
41	224	62.4	68	2	08NKE7_XANCP
42	224	62.4	75	2	05H1P7_XANOR
43	223	62.1	68	1	CSPE_BUCBP
44	223	62.1	68	2	07WDH6_BORBR
45	223	62.1	68	2	07W212_BORBP

ALIGNMENTS

RESULT 1

CSPA_MYCBO STANDARD; PRT; 67 AA.

AC P63843; 006360; 35, Created
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Probable cold shock protein A.
GN Name=csps; OrderedLocNames=Mb3672c;
OS Mycobacterium bovis.
OC Bacteria; Actinobacteria; Actinomycetales;
OC Corynebacteriaceae; Mycobacteriaceae; Mycobacterium;
OC Mycobacterium tuberculosis complex.
OX NCBI_TaxID=1765;
[1]

RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=AF2122/97; PubMed=12788972; DOI=10.1073/pnas.1130426100;
RX MEDLINE=22709107; PubMed=12788972; DOI=10.1073/pnas.1130426100;
RA Garnier T., Elgimeier K., Camus J.-C., Medina N., Mansoor H.,
RA Pryor M., Dutfoy S., Grondin S., Lacroix C., Monsemp C., Simon S.,
RA Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,
RA Parkhill J., Barrell B.G., Cole S.T., Gordon S.V., Hewison R.G.,
RT "The complete genome sequence of Mycobacterium bovis."
RL Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882(2003).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Contains 1 CSD (cold-shock) domain.

CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.

CC EMBL; BX248346; CAD95858.1; -; Genomic DNA.

DR HSSP; P41016; 1HZC.

DR InterPro; IPR012156; Cold_shock_Cspa.

DR InterPro; IPR011129; CSP.

DR InterPro; IPR002059; CSP_DNA_bd.

DR InterPro; IPR012340; OB_NA_bd_sud.

DR Pfam; PF00313; CSD; 1.

DR PIRSF; PIRSF02599; Cold_shock_A; 1.

DR PRINTS; PR00050; COLDSHOCK.

DR PRODOM; PD000621; Cold_shock; 1.

DR SMART; SM00357; CSP; 1; Cold_shock; 1.

DR PROSITE; PS00352; COLD_SHOCK; 1.

DR Activator; Complete proteome; DNA-binding; Transcription;

KW Transcription; Complete proteome; DNA-binding; Transcription;

FT DOMAIN 4 64 CSD.

SO SEQUENCE 67 AA; 7370 MW; F87506E09EB87C82 CRC64;

Query Match 100.0%; Score 359; DB 1; Length 67;
Best Local Similarity 100.0%; Pred. No. 8.2e-34;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MPQGTWKWNAEKGFPIAPEDGSADVFWYHYTEIOGTGRTLEBNKVFEIIGHSPKGPQ 60
DB 1 MPQGTWKWNAEKGFPIAPEDGSADVFWYHYTEIOGTGRTLEBNKVFEIIGHSPKGPQ 60
QY 61 ATGVRS 67
DB 61 ATGVRS 67

RESULT 2
CSPA_MYCTU STANDARD; PRT; 67 AA.
ID CSPA_MYCTU
AC P63848; O06360;
DT 01-NOV-1997 (Rel. 35, Created)
DT 25-OCT-2004 (Rel. 45, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Probable cold shock protein A.
GN Name=cspa; OrderedLocustNames=RV3648c, MT3750.1; ORFNames=MTCY15C10.04;
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacteriaceae; Mycobacteriaceae; Mycobacterium;
OC Mycobacterium tuberculosis complex.
OC NCBI_TaxId=1773;
RN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=H37Rv;
RX MEDLINE=98295967; PubMed=9634230; DOI=10.1038/31159;
RA Cole S.T., Broch R., Parkhill J., Garnier T., Church C.M.,
RA Harris D.E., Gordon S.V., Eiglmeier K., Gas S., Barry C.B. III,
RA Tekala F., Badcock K., Basham D., Brown D., Chillingworth T.,
RA Connor R., Davies R.M., Devlin K., Fellwell T., Gentles S., Hamlin N.,
RA Holroyd S., Hornsby T., Jagels K., Krogh A., McLean J., Moule S.,
RA Murphy L.D., Oliver S., Osborne J., Quail M.A., Rajandream M.A.,
RA Rogers J., Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrett B.G.
RT Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence.
RL Nature 393:537-544(1998).
RN 12]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=CDC 1551 / Oshkosh;
RX MEDLINE=22206494; PubMed=12218036;
RA DOI=10.1128/JB.184.19.5479-5490.2002;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J.D., Deboy R.T., Dodson R.J., Gwinn M.L., Haft D.H.,
RA Hickey R.K., Kolony J.F., Nelson W.C., Umayan L.A., Esmolaeva M.D.,
RA Salzberg S.L., Delcher A., Uitterback T.R., Weidman J.F., Khouri H.M.,
RA Gill J., Mikula A., Bishai W., Jacobs W.R. Jr., Venter J.C.,
RA Fraser C.M.;
RT "Whole-genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains."
RL J. Bacteriol. 184:5479-5490(2002).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Contains 1 CSD (cold-shock) domain.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL: BX842583; CAB08840.1; -; Genomic DNA.
CC EMBL: AB000516; AAK48111.1; -; Genomic DNA.
CC PIR: A70564; A70564.
CC HSSP: P41016; IHZC.
CC TIGR: MT3750.1; -.
CC TubercuList: RV3648c; -.
CC InterPro: IPR012156; Cold_shock_Cspa.
CC InterPro: IPR011129; CSP.
CC InterPro: IPR002059; CSP_DNA_bd.
CC InterPro: IPR012340; OB_NA_bd_sub.
CC Pfam: PF00313; CSD, 1.
CC PIRSF: PIRSF002599; Cold_shock_A; 1.

```

```

DR PRINTS: PR00050; COLDSHOCK.
DR ProDom: PD000621; Cold_shock; 1.
DR SMART: SM00357; CSP, 1.
DR PROSITE: PS00352; COLD_SHOCK; 1.
DR Activator: Complete proteome; DNA-binding; Transcription;
DR Transcription regulation.
DR DOMAIN 4 64 CSD.
DR SEQUENCE 67 AA; 7370 MW; F87506E09EB87C82 CRC64;

Query Match 100.0%; Score 359; DB 1; Length 67;
Best local Similarity 100.0%; Pred. No. 8,2e-34;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPQGTWKWNAEKGFPIAPEDGSADVFWYHYTEIOGTGRTLEBNKVFEIIGHSPKGPQ 60
DB 1 MPQGTWKWNAEKGFPIAPEDGSADVFWYHYTEIOGTGRTLEBNKVFEIIGHSPKGPQ 60
QY 61 ATGVRS 67
DB 61 ATGVRS 67

RESULT 3
0744B8_MYCPA PRELIMINARY; PRT; 67 AA.
ID 0744B8_MYCPA
AC 0744B8;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Cspa_1.
GN Name=cspa; OrderedLocustNames=MAP0423;
OS Mycobacterium paratuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacteriaceae; Mycobacteriaceae; Mycobacterium;
OC Mycobacterium avium complex (MAC).
OC NCBI_TaxId=1770;
RN 11]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=K10;
RL I. L., Bannantine J., Zhang Q., Amonsin A., Alt D., Kapur V.;
RA Submitted (SEP-2003) to the EMBL/Genbank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Contains 1 CSD (cold-shock) domain.
CC EMBL: AB017228; AA02740.1; -; Genomic DNA.
CC GO: GO:000677; rDNA binding; IEA.
CC GO: GO:006355; P-regulation of transcription; IEA.
CC GO: GO:006350; P-transcription; IEA.
DR InterPro: IPR002059; Cold_shock.
DR InterPro: IPR012156; Cold_shock_Cspa.
DR InterPro: IPR011129; CSP.
DR Pfam: PF00313; CSD, 1.
DR PIRSF: PIRSF002599; Cold_shock_A; 1.
DR PRINTS: PR000621; Cold_shock; 1.
DR ProDom: PD000621; Cold_shock; 1.
DR SMART: SM00357; CSP, 1.
DR PROSITE: PS00352; COLD_SHOCK; 1.
DR Activator: Complete proteome; DNA-binding; Transcription;
DR Transcription regulation.
DR SEQUENCE 67 AA; 7370 MW; F87506E09EB87C82 CRC64;

Query Match 100.0%; Score 359; DB 2; Length 67;
Best local Similarity 100.0%; Pred. No. 8,2e-34;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPQGTWKWNAEKGFPIAPEDGSADVFWYHYTEIOGTGRTLEBNKVFEIIGHSPKGPQ 60
DB 1 MPQGTWKWNAEKGFPIAPEDGSADVFWYHYTEIOGTGRTLEBNKVFEIIGHSPKGPQ 60
QY 61 ATGVRS 67
DB 61 ATGVRS 67

```

RESULT 4
ID 069550 MYCBL PRELIMINARY; PRT; 67 AA.
AC 069550;
DT 01-AUG-1998 (TREMblrel. 07, Created)
DT 01-AUG-1998 (TREMblrel. 07, Last sequence update)
DT 01-FEB-2005 (TREMblrel. 29, Last annotation update)
DE Small cold-shock protein (putative cold shock protein).
GN Name=cspA; Synonyms=cspA; OrderedLocNames=MD0198;
OS Mycobacterium leprae.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1769;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=93188700; PubMed=8446027;
RA Biglieri K., Honore N., Woods S.A., Caudron B., Cole S.T.;
RT "Use of an ordered cosmid library to deduce the genomic organization
of Mycobacterium leprae."
RL Mol. Microbiol. 7:197-206(1993).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Badcock K., Churcher C.M.;
RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RA Parkhill J., Barrell B.G., Rajandream M.A.;
RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
RN [4]
RP NUCLEOTIDE SEQUENCE.
RX STRAIN=TN;
RC MEDLINE=2118733; PubMed=11234002; DOI=10.1038/35059006;
RA Cole S.T., Biglieri K., Parkhill J., James K.D., Thomson N.R.,
RA Wheeler P.R., Honore N., Garner T., Churcher C.M., Harris D.E.,
RA Mungall K.L., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R.M., Devlin K., Duthoy S., Feltwell T., Fraser A., Hamlin N.,
RA Holroyd S., Hornby K., Jagsels K., Lacroix C., Maclean J., Moulé S.,
RA Murphy L.D., Oliver K., Quail M.A., Rajandream M.A., Rutherford K.M.,
RA Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R.,
RA Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,
RA Barrell B.G.;
RT "Massive gene decay in the leprosy bacillus."
RL Nature 409:1007-1011(2001).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Contains 1 CSD (cold-shock) domain.
DR EMBL: AL023093; CA118920.1; -; Genomic DNA.
DR EMBL: AL583917; CAC29706.1; -; Genomic DNA.
DR PIR: P86933; P86933.
DR HSP; P41016; IHZC.
DR Leproma; ML0198; -;
DR GO: GO:0003677; F:DNA binding; IEA.
DR GO: GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR GO: GO:0006350; P:transcription; IEA.
DR InterPro: IPR002059; Cold_shock.
DR InterPro: IPR012156; Cold_shock.
DR InterPro: IPR011129; CSP.
DR InterPro: IPR012340; NA-bind_OB_sub.
DR Pfam: PF00313; CSD; 1.
DR PIRSF: PIRSF002599; Cold_shock_A; 1.
DR PRINTS: PR00050; COLDSHOCK.
DR ProDom: PD000621; Cold_shock; 1.
DR SMART: SM00357; CSP; 1.
DR PROSITE: PS00352; COLD_SHOCK; 1.
KW Activator; Complete proteome; DNA-binding; Transcription;
KW Transcription regulation.
SQ SEQUENCE 67 AA; 7342 MW; A88506E09BE87BF2 CRC64;
Query Match 98.1%; Score 352; DB 2; Length 67;
Best Local Similarity 97.0%; Pred. No. 5,3e-33;
Matches 65; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 1 MPOGTVMFNAKGFGLIAPEDGSADVFWHTETIGTGFRLEENKVFEGHSPKGPQ 60
|||||
|||||

DB 1 MPOGTVMFNAKGFGLIAPEDGSADVFWHTETIGTGFRLEENKVFEGHSPKGPQ 60
QY 61 ATGVRSI 67
|||||
DB 61 ATGVRSV 67
RESULT 5
ID 09KGM0 MYCSM PRELIMINARY; PRT; 67 AA.
AC 09KGM0;
DT 01-OCT-2000 (TREMblrel. 15, Created)
DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
DT 01-OCT-2003 (TREMblrel. 25, Last annotation update)
DE Cold-shock protein CspA.
GN Name=cspA;
OS Mycobacterium smegmatis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1772;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Shires K.L., Poles R., Streyn L.;
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Contains 1 CSD (cold-shock) domain.
DR EMBL: AF281675; AAF89211.1; -; Genomic DNA.
DR HSP; P41016; IHZC.
DR GO: GO:0003677; F:DNA binding; IEA.
DR GO: GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR GO: GO:0006350; P:transcription; IEA.
DR InterPro: IPR002059; Cold_shock.
DR InterPro: IPR012156; Cold_shock.
DR InterPro: IPR011129; CSP.
DR InterPro: IPR012340; NA-bind_OB_sub.
DR Pfam: PF00313; CSD; 1.
DR PIRSF: PIRSF002599; Cold_shock_A; 1.
DR PRINTS: PR00050; COLDSHOCK.
DR ProDom: PD000621; Cold_shock; 1.
DR SMART: SM00357; CSP; 1.
DR PROSITE: PS00352; COLD_SHOCK; 1.
KW Activator; DNA-binding; Transcription; Transcription regulation.
SQ SEQUENCE 67 AA; 7347 MW; 56C116809EF746EA CRC64;
Query Match 95.0%; Score 341; DB 2; Length 67;
Best Local Similarity 92.5%; Pred. No. 1e-31;
Matches 62; Conservative 4; Mismatches 1; Indels 0; Gaps 0;
QY 1 MPOGTVMFNAKGFGLIAPEDGSADVFWHTETIGTGFRLEENKVFEGHSPKGPQ 60
|||||
DB 1 MPOGTVMFNAKGFGLIAPEDGSADVFWHTETIGTGFRLEENKVFEGHSPKGPQ 60
QY 61 ATGVRSI 67
|||||
DB 61 ATGVRSI 67
RESULT 6
ID 0522Y3 NOCPA PRELIMINARY; PRT; 68 AA.
AC 0522Y3;
DT 25-OCT-2004 (TREMblrel. 28, Created)
DT 25-OCT-2004 (TREMblrel. 28, Last sequence update)
DT 25-OCT-2004 (TREMblrel. 28, Last annotation update)
DE Putative cold shock protein.
GN OrderedLocNames=nfa360;
OS Nocardia farcinica.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Nocardiaceae; Nocardia.
OX NCBI_TaxID=37329;
RN [1]
RP NUCLEOTIDE SEQUENCE.
SQ STRAIN=IFM 10152;
QY 1 MPOGTVMFNAKGFGLIAPEDGSADVFWHTETIGTGFRLEENKVFEGHSPKGPQ 60
|||||
|||||

RX PubMed=15466710; DOI=10.1073/pnas.0406410101;
 RA Ishikawa J., Yamashita A., Mikami Y., Hoshino Y., Kurita H., Hotta K.,
 RA Shiba T., Hattori M.;
 RT "The complete genomic sequence of *Nocardia farcinica* IFM 10152.";
 CC Proc. Natl. Acad. Sci. U.S.A. 101:14925-14930(2004).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 DR EMBL; AF006618; BAD5208.1; -; genomic_DNA.
 DR GO; GO:0003677; P:DNA binding; IEA.
 DR GO; GO:0006355; P:regulation of transcription; IEA.
 DR GO; GO:0006350; P:transcription; IEA.
 DR InterPro; IPR002059; Cold_shock_CSPA.
 DR InterPro; IPR01129; Cold_shock_CSPA.
 DR InterPro; IPR01129; Cold_shock_CSPA.
 DR PIRSF; PIRSF002599; Cold_shock_A; 1.
 DR PRINTS; PR00050; COLDSHOCK.
 DR ProDom; PD000621; Cold_shock; 1.
 DR SMART; SM00357; CSP; 1.
 DR PROSITE; PS00352; COLD_SHOCK; 1.
 DR Activator; Complete proteome; DNA-binding; Transcription;
 KW Transcription regulation.
 SO SEQUENCE 68 AA; 7338 MW; F961B7CC51D05CE6 CRC64;

Query Match 88.9%; Score 319; DB 2; Length 68;
 Best Local Similarity 88.1%; Pred. No. 3.6e-29;
 Matches 59; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 MPOGTWKWNAKGFPGFIAPEDGSADVFTYHTIQTGRTLEENKVEFEIGHSKGPQ 60
 DB 1 MPOGTWKWNAKGFPGFIAPEDGSADVFTYHTIQTGRTLEENKVEFEIGHSKGPQ 60

OY 61 ATGVRSI 67
 DB 61 ATGVRSI 67

RESULT 7
 OGNUS3_CORDI PRELIMINARY; PRT; 67 AA.
 ID OGNUS3_CORDI PRELIMINARY; PRT; 67 AA.
 AC OGNUS3; PRT; 67 AA.
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DB Cold-shock protein.
 GN Name=CSPA; OrderedAccession=DIP0320;
 OS Corynebacterium diphtheriae.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacteriaceae; Corynebacteriaceae; Corynebacterium.
 OX NCBI_TaxID=1717;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=Blocky graves / NCTC 13129;
 RC MEDLINE=22365443; PubMed=14602910; DOI=10.1093/nar/9/kg874;
 RA Cerdeno-Tarraga A.-M., Estratou A., Dover L.G., Holden M.T.G.,
 RA Pallen M.J., Bentley S.D., Besra G.S., Churcher C.M., James K.D.,
 RA de Souza A., Chillingworth T., Cronin A., Dowd L., Felkwell T.,
 RA Hamlin N., Holtroyd S., Jagsels K., Moule S., Quail M.A.,
 RA Rabinovitch E., Rutherford K.M., Thomson N.R., Unwin L.,
 RA Whitehead S., Barrett B.G., Parkhill J.;
 RT "The complete genome sequence and analysis of *Corynebacterium*
 RT *diphtheriae* NCTC13129.";
 CC Nucleic Acids Res. 31:6516-6523(2003).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- SIMILARITY: Contains 1 CSD (cold-shock) domain.
 DR EMBL; BX248354; CAE48825.1; -; genomic_DNA.
 DR GO; GO:0003677; P:DNA binding; IEA.
 DR GO; GO:0006355; P:regulation of transcription; IEA.
 DR GO; GO:0006350; P:transcription; IEA.
 DR InterPro; IPR002059; Cold_shock_CSPA.
 DR InterPro; IPR012156; Cold_shock_CSPA.
 DR InterPro; IPR01129; CSP.
 DR Pfam; PF00313; CSD; 1.

DR PIRSF; PIRSF002599; Cold_shock_A; 1.
 DR PRINTS; PR00050; COLDSHOCK.
 DR ProDom; PD000621; Cold_shock; 1.
 DR SMART; SM00357; CSP; 1.
 DR PROSITE; PS00352; COLD_SHOCK; 1.
 DR Activator; Complete proteome; DNA-binding; Transcription;
 KW Transcription regulation.
 SO SEQUENCE 67 AA; 7347 MW; E837CC41CFACE3B CRC64;

Query Match 84.1%; Score 302; DB 2; Length 67;
 Best Local Similarity 83.6%; Pred. No. 3.3e-27;
 Matches 56; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 1 MPOGTWKWNAKGFPGFIAPEDGSADVFTYHTIQTGRTLEENKVEFEIGHSKGPQ 60
 DB 1 MPOGTWKWNAKGFPGFIAPEDGSADVFTYHTIQTGRTLEENKVEFEIGHSKGPQ 60

OY 61 ATGVRSI 67
 DB 61 AOVVRSI 67

RESULT 8
 Q4JSS1_CORJK PRELIMINARY; PRT; 67 AA.
 ID Q4JSS1_CORJK PRELIMINARY; PRT; 67 AA.
 AC Q4JSS1; PRT; 67 AA.
 DT 13-SEP-2005 (TrEMBLrel. 31, Created)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
 DB Putative cold shock protein.
 GN Name=CSPA; ORFNames=Jk1955;
 OS Corynebacterium jeikeium (strain K411).
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacteriaceae; Corynebacteriaceae; Corynebacterium.
 OX NCBI_TaxID=306537;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=K411;
 RC PubMed=15968079; DOI=10.1128/JB.187.13.4671-4682.2005;
 RA Tauch A., Kaiser O., Hain T., Goessmann A., Weisshaar B.,
 RA Albersmeier A., Bekel T., Bischoff N., Brune I., Chakraborty T.,
 RA Kallinowski J., Meyer F., Rupp O., Schneider S., Viehoveer P.,
 RA Fuhrer A.;
 RT "Complete Genome Sequence and Analysis of the Multiresistant
 RT Nosocomial Pathogen *Corynebacterium jeikeium* K411, a Lipid-Regulating
 RT Bacterium of the Human Skin Flora.";
 RL J. Bacteriol. 187:4671-4682(2005).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=K411;
 RC Linke B., Tauch A.;
 RL Submitted (DEC-2004) to the EMBL/Genbank/DBJ databases.
 DR EMBL; CR931997; CA138136.1; -; genomic_DNA.
 SO SEQUENCE 67 AA; 7267 MW; D60BAE57060DD3A3 CRC64;

Query Match 82.5%; Score 296; DB 2; Length 67;
 Best Local Similarity 82.1%; Pred. No. 1.7e-26;
 Matches 55; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 1 MPOGTWKWNAKGFPGFIAPEDGSADVFTYHTIQTGRTLEENKVEFEIGHSKGPQ 60
 DB 1 MPOGTWKWNAKGFPGFIAPEDGSADVFTYHTIQTGRTLEENKVEFEIGHSKGPQ 60

OY 61 ATGVRSI 67
 DB 61 AOVVRSI 67

RESULT 9
 O8FLY0_CORFP PRELIMINARY; PRT; 251 AA.
 ID O8FLY0_CORFP PRELIMINARY; PRT; 251 AA.
 AC O8FLY0; PRT; 251 AA.
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)


```
DT 01-MAR-2003 (TrEMBLrel. 23, last sequence update)
DR 01-OCT-2003 (TrEMBLrel. 25, last annotation update)
DB Hypothetical protein.
GN OrderedLocustNames=CR2729;
OS Corynebacterium efficiens.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacteriaceae; Corynebacteriaceae; Corynebacterium.
OX NCBI_TaxID=152794;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=YS-314 / AJ 12310 / DSM 44549 / JCM 11189;
RA MEDLINE=22723752; PubMed=12840036; DOI=10.1101/gr.1285603;
RA Niehio Y., Nakamura Y., Kawabayashi Y., Ueda Y., Kilmura B.,
RA Sugimoto S., Matsui K., Yamagishi A., Kikuchi H., Ikeo K.,
RA Gojohori T.;
RT "Comparative complete genome sequence analysis of the amino acid
RT replacements responsible for the thermostability of Corynebacterium
RT efficiens."
RL Genome Res. 13:1572-1579(2003).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Contains 1 CSD (cold-shock) domain.
DR EMBL; BA000035; BAC19539.1; -; Genomic_DNA.
DR HSSP; P32081; 1CSP.
DR GO; GO:0003677; P:DNA binding; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR InterPro; IPR002059; Cold_shock.
DR InterPro; IPR011129; CSP.
DR InterPro; IPR012340; NA-bind_OB_sub.
DR Pfam; PF00313; CSP; 1.
DR PRINTS; PR00050; COLDSHOCK.
DR ProDom; PD000621; Cold_shock; 1.
DR SMART; SM00357; CSP; 1.
DR PROSITE; PS00352; COLD_SHOCK; 1.
KW Activator; Complete proteome; DNA-binding; Hypothetical protein;
KW Transcription; Transcription regulation.
SQ SEQUENCE 251 AA; 27583 MW; 0CA3D167507D8451 CRC64;

Query Match 80.5%; Score 289; DB 2; Length 251;
Best Local Similarity 80.6%; Pred. No. 4,5e-25;
Matches 54; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

QY 1 MPOGTVMFAEKGFPIAPEDGSADVFNHTYETIGTGFTLEENQKVEFEIGHSPKQP 60
DB 185 MAQGTVMFAEKGFPIAPEDGSADVFNHTYETIGTGFTLEENQKVEFEIGHSPKQP 244
QY 61 ATGVRS 67
DB 245 AQQVRL 251

RESULT 10
Q8NTX7_CORGL PRELIMINARY; PRT; 67 AA.
ID Q8NTX7_CORGL PRELIMINARY; PRT; 67 AA.
AC Q8NTX7; Q6M813;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, last annotation update)
DE Cold shock proteins (COLD-SHOCK PROTEIN CSP).
GN Name=CSPA; OrderedLocustNames=Cg10174, Cg0215;
OS Corynebacterium glutamicum (Brevibacterium flavum).
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacteriaceae; Corynebacteriaceae; Corynebacterium.
OX NCBI_TaxID=1718;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;
RA Nakagawa S.;
RT "Complete genomic sequence of Corynebacterium glutamicum ATCC 13032.";
RT Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;
```

```
RX MEDLINE=22830012; PubMed=12948626; DOI=10.1016/S0168-1656(03)00154-8;
RA Kallinowski J., Busch N., Bartels D., Bischoff N., Bott M.,
RA Burkowski A., Busch N., Eggeling L., Eikmanns B.J., Galsgalat L.,
RA Goemann A., Hartmann M., Hutmacher K., Kraemer R., Linke B.,
RA McHardy A.C., Meyer F., Moeckel B., Pfeifferle W., Pehler A.,
RA Ray D.A., Rueckert C., Rupp O., Sahn H., Wendisch V.F., Wiegand I.,
RA Tsuch A.;
RT "The complete Corynebacterium glutamicum ATCC 13032 genome sequence
RT and its impact on the production of L-aspartate-derived amino acids
RT and vitamins."
RL J. Biotechnol. 104:5-25(2003).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Contains 1 CSD (cold-shock) domain.
DR EMBL; BA000036; BAB97567.1; -; Genomic_DNA.
DR EMBL; BX927148; CAP18741.1; -; Genomic_DNA.
DR HSSP; P32081; 1CSP.
DR GO; GO:0003677; P:DNA binding; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR InterPro; IPR002059; Cold_shock.
DR InterPro; IPR012156; Cold_shock_CSPA.
DR InterPro; IPR011129; CSP.
DR InterPro; IPR012340; NA-bind_OB_sub.
DR Pfam; PF00313; CSP; 1.
DR PRINTS; PR00050; COLD_SHOCK.
DR ProDom; PD000621; Cold_shock; 1.
DR SMART; SM00357; CSP; 1.
DR PROSITE; PS00352; COLD_SHOCK; 1.
KW Activator; Complete proteome; DNA-binding; Transcription;
KW Transcription regulation.
SQ SEQUENCE 67 AA; 7287 MW; C5FA343462BECB93 CRC64;

Query Match 78.6%; Score 282; DB 2; Length 67;
Best Local Similarity 77.6%; Pred. No. 7e-25;
Matches 52; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

QY 1 MPOGTVMFAEKGFPIAPEDGSADVFNHTYETIGTGFTLEENQKVEFEIGHSPKQP 60
DB 1 MAQGTVMFAEKGFPIAPEDGSADVFNHTYETIGTGFTLEENQKVEFEIGHSPKQP 60
QY 61 ATGVRS 67
DB 61 AQQVRL 67

RESULT 11
Q8NTX9_CORGL PRELIMINARY; PRT; 67 AA.
ID Q8NTX9_CORGL PRELIMINARY; PRT; 67 AA.
AC Q8NTX9; Q6M854;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, last annotation update)
DE Cold shock proteins (COLD-SHOCK PROTEIN CSP).
GN Name=CSPA2; OrderedLocustNames=Cg10308, Cg0371;
OS Corynebacterium glutamicum (Brevibacterium flavum).
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacteriaceae; Corynebacteriaceae; Corynebacterium.
OX NCBI_TaxID=1718;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;
RA Nakagawa S.;
RT "Complete genomic sequence of Corynebacterium glutamicum ATCC 13032.";
RT Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;
RA Kallinowski J., Busch N., Bartels D., Bischoff N., Bott M.,
RA Burkowski A., Busch N., Eggeling L., Eikmanns B.J., Galsgalat L.,
RA Goemann A., Hartmann M., Hutmacher K., Kraemer R., Linke B.,
RA McHardy A.C., Meyer F., Moeckel B., Pfeifferle W., Pehler A.,
```

RA Rey D.A., Rueckert C., Rupp O., Salm H., Wendisch V.F., Wiegand I.,
 RA Tach A.;
 RT "The complete *Corynebacterium glutamicum* ATCC 13032 genome sequence
 RT and its impact on the production of L-aspartate-derived amino acids
 RT and vitamins", 104:5-25(2003).
 RL J. Biotechnol. 104:5-25(2003).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- SIMILARITY: Contains 1 CSD (cold-shock) domain.
 DR EMBL: BA000036; BAB97701.1; -; Genomic DNA.
 DR EMBL: BX927148; CAP18878.1; -; Genomic DNA.
 DR HSSP: P32081; 1CSP.
 DR GO: GO:0003677; F:DNA binding; IEA.
 DR GO: GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
 DR GO: GO:0006350; P:transcription; IEA.
 DR InterPro: IPR002059; Cold shock.
 DR InterPro: IPR012156; Cold_shock_CspsA.
 DR InterPro: IPR011129; CSP.
 DR InterPro: IPR012340; NA-bind_OB_sub.
 DR Pfam: PF00313; CSD; 1.
 DR PIRSF: PIRSF002599; Cold_shock_A; 1.
 DR PRINTS: PR00050; COLDSHOCK.
 DR ProDom: PD000621; Cold_shock; 1.
 DR SMART: SM00357; CSP; 1.
 DR PROSITE: PS00352; COLD_SHOCK; 1.
 DR Activator; Complete proteome; DNA-binding; Transcription;
 KW Transcription regulation.
 SQ SEQUENCE 67 AA; 7290 MW; AF292D90C0E8ACD0 CRC64;

Query Match 76.3%; Score 275; DB 2; Length 67;
 Best Local Similarity 77.6%; Pred. No. 4,5e-24;
 Matches 52; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

QY 1 MPOGTYKMFNAKGFGEFIAPEDGSADVFAHYTEIQGTGRTLEENKVFEFGHSPKGPQ 60
 DB 1 MPOGTYKMFNAKGFGEFIAPEDGSADVFAHYTEIQGTGRTLEENKVFEFGHSPKGPQ 60

QY 61 ATGVRSL 67
 DB 61 AQAVRAI 67

RESULT 12
 Q8FS6 COREP PRELIMINARY; PRT; 67 AA.
 ID Q8FS6 COREP PRELIMINARY; PRT; 67 AA.
 AC Q8FS6;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
 DE Putative cold shock protein.
 DE OrderedLocNames=CSD0306;
 OS *Corynebacterium efficiens*.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC *Corynebacteriaceae*; *Corynebacterium*.
 NCBI_TaxID=152794;
 RX NUCLEOTIDE SEQUENCE.
 RP STRAIN=YS-314 / AJ 12310 / DSM 44549 / JCM 11189;
 RC MEDLINE=2273732; PubMed=12840036; DOI=10.1101/gr.1285603;
 RA Nishio Y., Nakamura Y., Kawarabayashi Y., Usuda Y., Kimura E.,
 RA Sugimoto S., Matsui K., Yamagishi A., Kikuchi H., Iseo K.,
 RA Gojohori T.;
 RT "Comparative complete genome sequence analysis of the amino acid
 RT replacements responsible for the thermostability of *Corynebacterium*
 RT *efficiens*."; 13:1572-1579(2003).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- SIMILARITY: Contains 1 CSD (cold-shock) domain.
 DR EMBL: BA000035; BAC17116.1; -; Genomic DNA.
 DR HSSP: P32081; 1CSP.
 DR GO: GO:0003677; F:DNA binding; IEA.
 DR GO: GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
 DR GO: GO:0006350; P:transcription; IEA.
 DR InterPro: IPR002059; Cold_shock.

DR InterPro: IPR012156; Cold_shock_CspsA.
 DR InterPro: IPR011129; CSP.
 DR InterPro: IPR012340; NA-bind_OB_sub.
 DR Pfam: PF00313; CSD; 1.
 DR PIRSF: PIRSF002599; Cold_shock_A; 1.
 DR PRINTS: PR00050; COLDSHOCK.
 DR ProDom: PD000621; Cold_shock; 1.
 DR SMART: SM00357; CSP; 1.
 DR PROSITE: PS00352; COLD_SHOCK; 1.
 DR Activator; Complete proteome; DNA-binding; Transcription;
 KW Transcription regulation.
 SQ SEQUENCE 67 AA; 7292 MW; BDB92D90D8A8CD0 CRC64;

Query Match 76.3%; Score 274; DB 2; Length 67;
 Best Local Similarity 77.6%; Pred. No. 5,9e-24;
 Matches 52; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

QY 1 MPOGTYKMFNAKGFGEFIAPEDGSADVFAHYTEIQGTGRTLEENKVFEFGHSPKGPQ 60
 DB 1 MPOGTYKMFNAKGFGEFIAPEDGSADVFAHYTEIQGTGRTLEENKVFEFGHSPKGPQ 60

QY 61 ATGVRSL 67
 DB 61 AQAVRAI 67

RESULT 13
 Q4NFR7 9MICC PRELIMINARY; PRT; 67 AA.
 ID Q4NFR7 9MICC PRELIMINARY; PRT; 67 AA.
 AC Q4NFR7;
 DT 13-SEP-2005 (TrEMBLrel. 31, Created)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
 DE Cold-shock protein, DNA-binding.
 DE ORFNames=ArthDRAP1862;
 OS *Arthrobacter* sp. FB24.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC *Micrococcineae*; *Micrococcaceae*; *Arthrobacter*.
 NCBI_TaxID=290399;
 RX NUCLEOTIDE SEQUENCE.
 RP STRAIN=FB24;
 RC US DOE Joint Genome Institute (JGI-PGF);
 RA Copeland A., Lucas S., Lapidus A., Barry K., Detter C., Glavina T.,
 RA Hammon N., Israni S., Pitluck S., Richardson P.;
 RT "Sequencing of the draft genome assembly of *Arthrobacter* sp. FB24.";
 RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=FB24;
 RG US DOE Joint Genome Institute (PGF-ORNL);
 RA Lartimer F., Land M.;
 RT Annotation of the draft genome assembly of *Arthrobacter* sp. FB24.";
 RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
 CC -1- CAUTION: The sequence shown here is derived from an
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 CC preliminary data.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- SIMILARITY: Contains 1 CSD (cold-shock) domain.
 DR EMBL: AAHG01000006; BAI96230.1; -; Genomic DNA.
 DR InterPro: IPR012156; Cold_shock_CspsA.
 DR InterPro: IPR011129; CSP.
 DR InterPro: IPR002059; CSP_DNA_bd.
 DR Pfam: PF00313; CSD; 1.
 DR PIRSF: PIRSF002599; Cold_shock_A; 1.
 DR PRINTS: PR00050; COLDSHOCK.
 DR ProDom: PD000621; Cold_shock; 1.
 DR SMART: SM00357; CSP; 1.
 DR PROSITE: PS00352; COLD_SHOCK; 1.
 DR Activator; DNA-binding; Transcription; Transcription regulation.
 KW Transcription regulation.
 SQ SEQUENCE 67 AA; 7267 MW; 850595CD0BC77CF4 CRC64;

Query Match 74.1%; Score 266; DB 2; Length 67;

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using BW model

Run on: April 14, 2006, 17:32:59 ; Search time 6.19531 Seconds
(without alignments)
894.107 Million cell updates/sec

Title: US-10-620-246-147

Perfect score: 359
Sequence: 1 MPQGYKWFNAEKFGFIAP.....VBEIGHSPKQATGVSL 67

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/ptodata/1/1aa/5 COMB.pep:*
- 2: /cgn2_6/ptodata/1/1aa/6 COMB.pep:*
- 3: /cgn2_6/ptodata/1/1aa/H COMB.pep:*
- 4: /cgn2_6/ptodata/1/1aa/PCTUS COMB.pep:*
- 5: /cgn2_6/ptodata/1/1aa/RB COMB.pep:*
- 6: /cgn2_6/ptodata/1/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	359	100.0	67	2	US-09-050-739-147
2	282	78.6	67	2	US-09-603-208A-82
3	282	78.6	67	2	US-09-603-208A-84
4	275	76.6	67	2	US-09-603-208A-88
5	227.5	63.4	88	2	US-09-134-000C-4171
6	222	61.8	131	2	US-09-489-039A-13476
7	221	60.9	70	2	US-09-516-667-80
8	218.5	60.9	66	1	US-08-203-806B-11
9	218.5	60.9	66	2	US-09-017-754A-11
10	217	60.4	70	2	US-09-516-667-79
11	216	60.2	67	2	US-09-902-540-12349
12	216	60.2	69	1	US-08-203-806B-4
13	216	60.2	69	1	US-08-203-806B-9
14	216	60.2	69	2	US-09-017-754A-9
15	216	60.2	73	2	US-09-489-039A-14126
16	216	60.2	73	2	US-09-543-681A-4308
17	214	59.6	73	2	US-09-107-532A-6235
18	212.5	59.2	73	2	US-08-203-806B-7
19	211	58.8	70	1	US-08-017-754A-20
20	211	58.8	70	2	US-09-516-667-82
21	211	58.8	70	2	US-09-489-039A-10823
22	211	58.8	72	2	US-09-543-681A-5351
23	211	58.8	72	2	US-09-107-532A-4506
24	211	58.8	72	2	US-09-581-005-4
25	208.5	58.1	66	2	US-09-581-005-4
26	208.5	58.1	66	2	US-09-581-005-4
27	207.5	57.8	66	2	US-09-581-005-4

28	205	57.1	67	2	US-09-902-540-16294	Sequence 16294, A
29	205	57.1	70	2	US-09-293-427-2	Sequence 2, Appl
30	205	57.1	85	2	US-09-540-236-2549	Sequence 2549, Ap
31	204	56.8	68	2	US-09-902-540-16803	Sequence 16803, A
32	201.5	56.1	66	2	US-09-000-476-2	Sequence 2, Appl
33	201	56.0	82	2	US-09-543-681A-7943	Sequence 7943, Ap
34	200.5	55.8	79	2	US-09-134-001C-3081	Sequence 3081, Ap
35	199.5	55.6	67	1	US-08-203-806B-12	Sequence 12, Appl
36	199.5	55.6	71	2	US-09-017-754A-12	Sequence 12, Appl
37	199	55.4	71	1	US-08-203-806B-8	Sequence 8, Appl
38	199	55.4	71	1	US-08-203-806B-8	Sequence 8, Appl
39	199	55.4	71	2	US-09-017-754A-2	Sequence 2, Appl
40	199	55.4	71	2	US-09-017-754A-8	Sequence 8, Appl
41	199	55.4	71	2	US-09-516-667-81	Sequence 81, Appl
42	199	55.4	71	2	US-09-293-427-4	Sequence 4, Appl
43	199	55.4	71	2	US-09-492-709A-372	Sequence 372, Ap
44	198.5	55.3	66	2	US-09-902-540-14444	Sequence 14444, A
45	196.5	54.7	70	2	US-09-134-000C-5656	Sequence 5656, Ap

ALIGNMENTS

RESULT 1
US-09-050-739-147
; Sequence 147, Application US/09050739
; Patent No. 6641814
; GENERAL INFORMATION:
; APPLICANT: ANDERSEN, Peter
; APPLICANT: NIELSEN, Rikke
; APPLICANT: OESTINGER, Thomas
; APPLICANT: RASMUSSEN, Peter Blak
; APPLICANT: ROSENKRANDS, Ida
; APPLICANT: WELDMING, Karin
; APPLICANT: FLORIO, Walter
; TITLE OF INVENTION: NUCLEIC ACIDS FRAGMENTS AND POLYPEPTIDE FRAGMENTS
; FILE REFERENCE: 670001-2002.1
; CURRENT APPLICATION NUMBER: US/09/050,739
; CURRENT FILING DATE: 1998-03-30
; EARLIER APPLICATION NUMBER: 0376/97
; EARLIER FILING DATE: 1997-04-02
; EARLIER APPLICATION NUMBER: 1277/97
; EARLIER FILING DATE: 1997-11-10
; EARLIER APPLICATION NUMBER: 60/044,624
; EARLIER FILING DATE: 1997-04-18
; EARLIER APPLICATION NUMBER: 60/070,488
; EARLIER APPLICATION NUMBER: 60/070,488
; NUMBER OF SEQ ID NOS: 173
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 147
; LENGTH: 67
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-09-050-739-147

Query Match 100.0%; Pred 359; DB 2; Length 67;
Best Local Similarity 100.0%; Pred. No. 3.5e-40; Indels 0; Gaps 0;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 1 MPQGYKWFNAEKFGFIAPEDGSADVFFVYITEIOGTGFTLENOKVBEIGHSPKQ 60
Db 1 MPQGYKWFNAEKFGFIAPEDGSADVFFVYITEIOGTGFTLENOKVBEIGHSPKQ 60

CY 61 ATGVSL 67
Db 61 ATGVSL 67

RESULT 2
US-09-603-208A-82
; Sequence 82, Application US/09603208A
; Patent No. 6822084

```
/ GENERAL INFORMATION:
/ APPLICANT: Pompeius, Markus
/ APPLICANT: Kroeger, Burkhard
/ APPLICANT: Schroder, Hartwig
/ APPLICANT: Zelder, Oskar
/ APPLICANT: Habernauer, Gregor
/ APPLICANT: Lee, Heung-Shick
/ APPLICANT: Kim, Hyung-Joon
/ TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING STRESS,
/ TITLE OF INVENTION: RESISTANCE AND TOLERANCE PROTEINS
/ FILE REFERENCE: BGI-124CP
/ CURRENT APPLICATION NUMBER: US/09/603,208A
/ PRIOR FILING DATE: 2000-06-23
/ PRIOR APPLICATION NUMBER: 60/141031
/ PRIOR FILING DATE: 1999-06-25
/ PRIOR APPLICATION NUMBER: 60/142692
/ PRIOR FILING DATE: 1999-07-01
/ PRIOR APPLICATION NUMBER: 60/151214
/ PRIOR FILING DATE: 1999-08-27
/ PRIOR APPLICATION NUMBER: DE 19930429.7
/ PRIOR FILING DATE: 1999-07-09
/ PRIOR APPLICATION NUMBER: DE 19932914.1
/ PRIOR FILING DATE: 1999-07-14
/ PRIOR APPLICATION NUMBER: DE 19940764.9
/ PRIOR FILING DATE: 1999-08-27
/ PRIOR APPLICATION NUMBER: DE 19941382.7
/ PRIOR FILING DATE: 1999-08-31
/ NUMBER OF SEQ ID NOS: 306
/ LENGTH: 67
/ TYPE: PR1
/ ORGANISM: Corynebacterium glutamicum
/ US-09-603-208A-82

Query Match          78.6%; Score 282; DB 2; Length 67;
Best Local Similarity 77.6%; Pred. No. 5, 8e-30;
Matches 52; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

QY 1 MPOGTVMFNAXKGFPIAPEDGSADVFTYHTBIQGTGRTLEBNOKVEFEIGHSPKQP 60
    |||||
DB 1 MAGGTVMFNKGKGFPIAPNDGSADLFVHYSTIQSGGFRTLEBNQPVFEVEGAKGPFQ 60
    |||||

QY 61 ATGVRSI 67
    |||||
DB 61 AAGVRAL 67
    |||||

RESULT 3
US-09-603-208A-84
/ Sequence 84, Application US/09603208A
/ GENERAL INFORMATION:
/ APPLICANT: Pompeius, Markus
/ APPLICANT: Kroeger, Burkhard
/ APPLICANT: Schroder, Hartwig
/ APPLICANT: Zelder, Oskar
/ APPLICANT: Habernauer, Gregor
/ APPLICANT: Lee, Heung-Shick
/ APPLICANT: Kim, Hyung-Joon
/ TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING STRESS,
/ TITLE OF INVENTION: RESISTANCE AND TOLERANCE PROTEINS
/ FILE REFERENCE: BGI-124CP
/ CURRENT APPLICATION NUMBER: US/09/603,208A
```

```
/ CURRENT FILING DATE: 2000-06-23
/ PRIOR APPLICATION NUMBER: 60/141031
/ PRIOR FILING DATE: 1999-06-25
/ PRIOR APPLICATION NUMBER: 60/142692
/ PRIOR FILING DATE: 1999-07-01
/ PRIOR APPLICATION NUMBER: 60/151214
/ PRIOR FILING DATE: 1999-08-27
/ PRIOR APPLICATION NUMBER: DE 19930429.7
/ PRIOR FILING DATE: 1999-07-09
/ PRIOR APPLICATION NUMBER: DE 19932914.1
/ PRIOR FILING DATE: 1999-07-14
/ PRIOR APPLICATION NUMBER: DE 19940764.9
/ PRIOR FILING DATE: 1999-08-27
/ PRIOR APPLICATION NUMBER: DE 19941382.7
/ PRIOR FILING DATE: 1999-08-31
/ NUMBER OF SEQ ID NOS: 306
/ LENGTH: 67
/ TYPE: PR1
/ ORGANISM: Corynebacterium glutamicum
/ US-09-603-208A-84

Query Match          78.6%; Score 282; DB 2; Length 67;
Best Local Similarity 77.6%; Pred. No. 5, 8e-30;
Matches 52; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

QY 1 MPOGTVMFNAXKGFPIAPEDGSADVFTYHTBIQGTGRTLEBNOKVEFEIGHSPKQP 60
    |||||
DB 1 MAGGTVMFNKGKGFPIAPNDGSADLFVHYSTIQSGGFRTLEBNQPVFEVEGAKGPFQ 60
    |||||

QY 61 ATGVRSI 67
    |||||
DB 61 AAGVRAL 67
    |||||

RESULT 4
US-09-603-208A-88
/ Sequence 88, Application US/09603208A
/ GENERAL INFORMATION:
/ APPLICANT: Pompeius, Markus
/ APPLICANT: Kroeger, Burkhard
/ APPLICANT: Schroder, Hartwig
/ APPLICANT: Zelder, Oskar
/ APPLICANT: Habernauer, Gregor
/ APPLICANT: Lee, Heung-Shick
/ APPLICANT: Kim, Hyung-Joon
/ TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING STRESS,
/ TITLE OF INVENTION: RESISTANCE AND TOLERANCE PROTEINS
/ FILE REFERENCE: BGI-124CP
/ CURRENT APPLICATION NUMBER: US/09/603,208A
/ PRIOR FILING DATE: 2000-06-23
/ PRIOR APPLICATION NUMBER: 60/141031
/ PRIOR FILING DATE: 1999-06-25
/ PRIOR APPLICATION NUMBER: 60/142692
/ PRIOR FILING DATE: 1999-07-01
/ PRIOR APPLICATION NUMBER: 60/151214
/ PRIOR FILING DATE: 1999-08-27
/ PRIOR APPLICATION NUMBER: DE 19930429.7
/ PRIOR FILING DATE: 1999-07-09
/ PRIOR APPLICATION NUMBER: DE 19932914.1
/ PRIOR FILING DATE: 1999-07-14
/ PRIOR APPLICATION NUMBER: DE 19940764.9
/ PRIOR FILING DATE: 1999-08-27
/ PRIOR APPLICATION NUMBER: DE 19941382.7
/ PRIOR FILING DATE: 1999-08-31
```

```
/ PRIOR FILING DATE: 1999-07-08
/ PRIOR APPLICATION NUMBER: DE 19931541.8
/ PRIOR FILING DATE: 1999-07-08
/ PRIOR APPLICATION NUMBER: DE 19932209.0
/ PRIOR FILING DATE: 1999-07-09
/ PRIOR APPLICATION NUMBER: DE 19932230.9
/ PRIOR FILING DATE: 1999-07-09
/ PRIOR APPLICATION NUMBER: DE 19932914.1
/ PRIOR FILING DATE: 1999-07-14
/ PRIOR APPLICATION NUMBER: DE 19940764.9
/ PRIOR FILING DATE: 1999-08-27
/ PRIOR APPLICATION NUMBER: DE 19941382.7
/ PRIOR FILING DATE: 1999-08-31
/ NUMBER OF SEQ ID NOS: 306
/ SEQ ID NO 88
/ LENGTH: 67
/ TYPE: PRT
/ ORGANISM: Corynebacterium glutamicum
US-09-603-208A-88
```

```
Query Match          76.6%; Score 275; DB 2; Length 67;
Best Local Similarity 77.6%; Pred. No. 4,9e-29;
Matches 52; Conservative 4; Mismatches 11; Indels 0; Gaps 0;
```

```
QY 1 MPOGTVKWFNAEKGFPIAPEDGSADVFWHTYETIGTGFTLEENQKVFEEIGHSPKGPQ 60
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 1 MPOGTVKWFNAEKGFPIAPEDGSADVFWHTYETIGTGFTLEENQKVFEEIGHSPKGPQ 60
```

```
QY 61 ATGVRSI 67
    |||||:
DB 61 AQAARAI 67
```

```
RESULT 5
US-09-134-000C-4171
/ Sequence 4171, Application US/09134000C
/ Patent No. 6617156
/ GENERAL INFORMATION:
```

```
/ APPLICANT: Lynn Doucette-Stamm et al
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
/ TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
/ FILE REFERENCE: 032796-032
/ CURRENT APPLICATION NUMBER: US/09/134,000C
/ PRIOR FILING DATE: 1998-08-13
/ PRIOR APPLICATION NUMBER: US 60/055,778
/ PRIOR FILING DATE: 1997-08-15
/ NUMBER OF SEQ ID NOS: 6812
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 4171
/ LENGTH: 88
/ TYPE: PRT
/ ORGANISM: Enterococcus faecalis
US-09-134-000C-4171
```

```
Query Match          63.4%; Score 227.5; DB 2; Length 88;
Best Local Similarity 68.8%; Pred. No. 1,4e-22;
Matches 44; Conservative 7; Mismatches 12; Indels 1; Gaps 1;
```

```
QY 1 MPOGTVKWFNAEKGFPIAPEDGSADVFWHTYETIGTGFTLEENQKVFEEIGHSPKGPQ 60
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 23 MEGGTVKWFNAEKGFPIAPEDGSADVFWHTYETIGTGFTLEENQKVFEEIGHSPKGPQ 81
```

```
QY 61 ATGV 64
    |||||:
DB 82 AAVV 85
```

```
RESULT 6
US-09-489-039A-13476
/ Sequence 13476, Application US/09489039A
/ Patent No. 6610836
/ GENERAL INFORMATION:
/ APPLICANT: Gary Breton et. al
```

```
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
/ TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
/ FILE REFERENCE: 2709.2004001
/ CURRENT APPLICATION NUMBER: US/09/489,039A
/ PRIOR FILING DATE: 2000-01-27
/ PRIOR APPLICATION NUMBER: US 60/117,747
/ PRIOR FILING DATE: 1999-01-29
/ NUMBER OF SEQ ID NOS: 14342
/ SEQ ID NO 13476
/ LENGTH: 131
/ TYPE: PRT
/ ORGANISM: Klebsiella pneumoniae
US-09-489-039A-13476
```

```
Query Match          61.8%; Score 222; DB 2; Length 131;
Best Local Similarity 63.1%; Pred. No. 1,3e-21;
Matches 41; Conservative 7; Mismatches 17; Indels 0; Gaps 0;
```

```
QY 3 QGTVKWFNAEKGFPIAPEDGSADVFWHTYETIGTGFTLEENQKVFEEIGHSPKGPQAT 62
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 67 KGVKWFNAEKGFPIAPEDGSADVFWHTYETIGTGFTLEENQKVFEEIGHSPKGPQAT 126
```

```
QY 63 GYRSI 67
    |||||:
DB 127 NVIAL 131
```

```
RESULT 7
US-09-516-667-80
/ Sequence 80, Application US/09516667
/ Patent No. 6610533
/ GENERAL INFORMATION:
```

```
/ APPLICANT: Inouye, Masayori
/ APPLICANT: Wang, Nan
/ APPLICANT: Yamana, Kuniohshi
/ TITLE OF INVENTION: COLD-SHOCK REGULATORY ELEMENTS, CONSTRUCTS THEREOF, AND
/ TITLE OF INVENTION: METHODS OF USE
/ FILE REFERENCE: 1053-00
/ CURRENT APPLICATION NUMBER: US/09/516,667
/ PRIOR FILING DATE: 2001-08-01
/ NUMBER OF SEQ ID NOS: 87
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 80
/ LENGTH: 70
/ TYPE: PRT
/ ORGANISM: E. coli
US-09-516-667-80
```

```
Query Match          61.6%; Score 221; DB 2; Length 70;
Best Local Similarity 64.1%; Pred. No. 7,7e-22;
Matches 41; Conservative 6; Mismatches 17; Indels 0; Gaps 0;
```

```
QY 4 GTVKWFNAEKGFPIAPEDGSADVFWHTYETIGTGFTLEENQKVFEEIGHSPKGPQATG 63
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 7 GTVKWFNAEKGFPIAPEDGSADVFWHTYETIGTGFTLEENQKVFEEIGHSPKGPQATG 66
```

```
QY 64 VRSI 67
    |||||:
DB 67 VVTL 70
```

```
RESULT 8
US-08-203-806B-11
/ Sequence 11, Application US/08203806B
/ Patent No. 5714575
/ GENERAL INFORMATION:
/ APPLICANT: Inouye, Masayori
/ APPLICANT: Jones, Pamela
/ APPLICANT: Bichagaray, Jean-Pierre
/ APPLICANT: Weining, Jian
/ APPLICANT: Pollitt, N. Stephen
/ APPLICANT: Goldstein, Joel
/ TITLE OF INVENTION: Nucleic Acid Sequences, Stress-induced
```

TITLE OF INVENTION: Proteins and Uses Thereof
 NUMBER OF SEQUENCES: 18
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Welser & Associates
 STREET: 230 South Fifteenth Street, Suite 5000
 CITY: Philadelphia
 STATE: PA
 COUNTRY: U.S.A.
 ZIP: 19102
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.2
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/203,806B
 FILING DATE: 01-MAR-1994
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Welser, Gerard J.
 REGISTRATION NUMBER: 19,763
 REFERENCE/DOCKET NUMBER: 377.5998P
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 215-875-8383
 TELEFAX: 215-875-8394
 TELEX: 834809 WEISTAK
 INFORMATION FOR SEQ ID NO: 11:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 66 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 ORIGINAL SOURCE:
 ORGANISM: Streptomyces clavuligerus

Query Match	60.9%	Score 218.5	DB 1	Length 66
Best Local Similarity	65.6%	Pred. No. 1.5e-21		
Matches	42	Conservative	6	Mismatches 15; Indels 1; Gaps 1
Qy	1	MPGGVKKFNNAEKGKGFAPEDGSGADVFWHYTEIOGTGFRTELEBNOKYFRIHGSPPKPGQ		60
		:		:: :
Dd	1	MATGVKRNKNFAEKSGFTAAODGGPDVFVHSAINATGRSLSEENQVNNEFDVTHG-EGPQ		59
Qy				
	61	ATGV	64	
Dd	60	AENV	63	

RESULT 9
 US-09-017-754A-11
 Sequence 11, Application US/09017754A
 Patent No. 633191
 GENERAL INFORMATION:
 APPLICANT: INOUYE, MASAYORI
 JONES, PAMELA
 ETCHEGARAY, JEAN-PIERRE
 WEINING, JIANG
 POLLITT, N. STEPHEN
 GOLDSTEIN, JOEL
 TITLE OF INVENTION: NUCLEIC ACID SEQUENCES, STRESS-INDUCED
 PROTEINS AND USES THEREOF
 NUMBER OF SEQUENCES: 20
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: SCHRADER, HARRISON, SEGAL & LEWIS, LLP
 STREET: 1600 Market Street, Suite 3600
 CITY: Philadelphia
 STATE: PA
 COUNTRY: USA
 ZIP: 19103
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible

```

1 OPERATING SYSTEM: PC-DOS/MS-DOS
2 SOFTWARE: Patent In Release #1.0, Version #1.3G
3 CURRENT APPLICATION DATA:
4 APPLICATION NUMBER: US/09/017,754A
5 FILING DATE: 03-Feb-1998
6 ATTORNEY/AGENT INFORMATION:
7 NAME: Welser, Gerard J.
8 REGISTRATION NUMBER: 19,763
9 REFERENCE/DOCKET NUMBER: 1113-cip2-div-00
10 TELECOMMUNICATION INFORMATION:
11 TELEPHONE: 215-751-2427
12 TELEFAX: 215-751-2658
13 INFORMATION FOR SEQ ID NO: 11:
14 SEQUENCE CHARACTERISTICS:
15 LENGTH: 66 amino acids
16 TYPE: amino acid
17 STRANDEDNESS: <Unknown>
18 TOPOLOGY: linear
19 MOLECULE TYPE: protein
20 SEQUENCE DESCRIPTION: SEQ ID NO: 11:
21 US-09-017-754A-11

```

```

Query Match      60.9%; Score 218.5; DB 2; Length 66;
Beet Local Similarity 65.6%; Pred. No. 1.5e-21;
Matches 42; Conservative 6; Mismatches 15; Indels 1; Gaps 1;

Oy 1 MPOGTVMFNNAKGGFRIAPEDGSADVFMHTSIQGTGFTLEENOKVEFEGHSPKQPO 60
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 MATVTVMFNNAKGGFRIADGGGPDVFMHTSAINATGFRLEENOVNVFDTVG-EGPO 59
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

Oy 61 ATGV 64
    |||||
Db 60 AENV 63
    |||||

```

```

RESULT 10
US-09-516-667-79
: Sequence 79, Application US/09516667
: Patent No. 6610533
: GENERAL INFORMATION:
: APPLICANT: Inouye, Masayori
: APPLICANT: Wang, Nan
: APPLICANT: Yamataka, Kunitoshi
: TITLE OF INVENTION: COLD-SHOCK REGULATORY ELEMENTS, CONSTRUCTS THEREOF, AND
: TITLE OF INVENTION: METHODS OF USE
: FILE REFERENCE: 1053-00
: CURRENT APPLICATION NUMBER: US/09/516,667
: CURRENT FILING DATE: 2001-08-01
: NUMBER OF SEQ. ID NOS: 87
: SOFTWARE: PatentIn Ver. 2.1
: SEQ. ID NO. 79
: LENGTH: 70
: TYPE: PRT
: ORGANISM: E. coli
: US-09-516-667-79

```

[illegible]

RESULT 11
US-09-902-540-12349
; Sequence 12349, Application US/09902540
; Patent No. 6833447


```

; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(11849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; PRIOR APPLICATION NUMBER: 2001-07-10
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ. ID NOS: 16825
; SEQ ID NO 12349
; LENGTH: 67
; TYPE: PR1
; ORGANISM: Myxococcus xanthus
; US-09-902-540-12349

Query Match      60.2%; Score 216; DB 2; Length 67;
Best Local Similarity 62.1%; Pred. No. 3.3e-21;
Matches 41; Conservative 6; Mismatches 19; Indels 0; Gaps 0;

QY      1 MPGTVKMFNAKGGFPIAPEDGSADVFWHTTETGCTGFRLEENQKVEFIQHSPPKQPT 60
DB      1 MATGTVMFNDAGFGFTVODGGEDVFWHSALNMDGFRITLQEGQKVEFVGKPGKLQ 60

QY      61 ATGVRG 66
DB      61 ANTVRA 66

RESULT 12
US-08-203-806B-4
; Sequence 4, Application US/08203806B
; Patent No. 5714575
; GENERAL INFORMATION:
; APPLICANT: Inouye, Masayori
; APPLICANT: Jones, Pamela
; APPLICANT: Etchegaray, Jean-Pierre
; APPLICANT: Weinig, Jian
; APPLICANT: Pollitt, N. Stephen
; APPLICANT: Goldstein, Joel
; TITLE OF INVENTION: Nucleic Acid Sequences, Stresses-Induced
; TITLE OF INVENTION: Proteins and Uses Thereof
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSES: Weiser & Associates
; STREET: 230 South Fifteenth Street, Suite 500
; CITY: Philadelphia
; STATE: PA
; COUNTRY: U.S.A.
; ZIP: 19102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Releasee #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/203,806B
; FILING DATE: 01-MAR-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Weiser, Gerard J.
; REGISTRATION NUMBER: 19,763
; REFERENCE/DOCKET NUMBER: 377.5998P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-875-8394
; TELEFAX: 215-875-8394
; TELEX: 834809 WEISYAK
; INFORMATION FOR SEQ. ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 69 amino acids
; TYPE: amino acid
```

```

; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-203-806B-4

Query Match      60.2%; Score 216; DB 1; Length 69;
Best Local Similarity 61.5%; Pred. No. 3.5e-21;
Matches 40; Conservative 6; Mismatches 19; Indels 0; Gaps 0;

QY      3 OGTVKMFNAKGGFPIAPEDGSADVFWHTTETGCTGFRLEENQKVEFIQHSPPKQPT 62
DB      5 KGQVKNFESKGGFTTPADGSDVFWHSALQNGFKTLAEQNVFEFIQDQKGPAAV 64

QY      63 GVRSL 67
DB      65 NVTAI 69

RESULT 13
US-08-203-806B-9
; Sequence 9, Application US/08203806B
; Patent No. 5714575
; GENERAL INFORMATION:
; APPLICANT: Inouye, Masayori
; APPLICANT: Jones, Pamela
; APPLICANT: Etchegaray, Jean-Pierre
; APPLICANT: Weinig, Jian
; APPLICANT: Pollitt, N. Stephen
; APPLICANT: Goldstein, Joel
; TITLE OF INVENTION: Nucleic Acid Sequences, Stresses-Induced
; TITLE OF INVENTION: Proteins and Uses Thereof
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSES: Weiser & Associates
; STREET: 230 South Fifteenth Street, Suite 500
; CITY: Philadelphia
; STATE: PA
; COUNTRY: U.S.A.
; ZIP: 19102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Releasee #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/203,806B
; FILING DATE: 01-MAR-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Weiser, Gerard J.
; REGISTRATION NUMBER: 19,763
; REFERENCE/DOCKET NUMBER: 377.5998P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-875-8394
; TELEFAX: 215-875-8394
; TELEX: 834809 WEISYAK
; INFORMATION FOR SEQ. ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 69 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Escherichia coli
; US-08-203-806B-9

Query Match      60.2%; Score 216; DB 1; Length 69;
Best Local Similarity 61.5%; Pred. No. 3.5e-21;
Matches 40; Conservative 6; Mismatches 19; Indels 0; Gaps 0;

QY      3 OGTVKMFNAKGGFPIAPEDGSADVFWHTTETGCTGFRLEENQKVEFIQHSPPKQPT 62
DB      5 KGQVKNFESKGGFTTPADGSDVFWHSALQNGFKTLAEQNVFEFIQDQKGPAAV 64
```


GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 14, 2006, 18:27:13 ; Search time 23.512 Seconds
(without alignments)
1190.652 Million cell updates/sec

Title: US-10-620-246-147

Perfect score: 359

Sequence: 1 MPOGTVKMFNAEKGFPIADP.....VPEFIGHSPKQATGVRSI 67

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	359	100.0	67	3	US-09-791-171-147
2	359	100.0	67	3	US-09-804-380-147
3	359	100.0	67	4	US-10-620-246-147
4	282	78.6	67	3	US-09-738-626-3695
5	282	78.6	67	5	US-10-953-856-51
6	275	76.6	67	3	US-09-738-626-3845
7	256	71.3	67	4	US-10-282-122A-62113
8	235	65.5	67	4	US-10-156-761-12354
9	230.5	64.2	68	4	US-10-156-761-11466
10	223	62.1	70	5	US-10-953-856-19
11	221	61.6	67	4	US-10-156-761-8364
12	216	60.2	69	5	US-10-953-856-13
13	213.5	59.5	66	4	US-10-282-122A-57331
14	211	58.8	70	3	US-09-912-020-328
15	211	58.8	70	4	US-10-282-122A-42639
16	211	58.8	70	4	US-10-282-122A-59825
17	211	58.8	70	4	US-10-282-122A-73080
18	211	58.8	70	4	US-10-282-122A-75306
19	211	58.8	70	4	US-10-282-122A-75459
20	211	58.8	70	5	US-10-771-241-328
21	211	58.8	70	5	US-10-953-856-1
22	211	58.8	70	5	US-10-953-856-63
23	211	58.8	70	5	US-10-953-856-93
24	211	58.8	70	6	US-11-067-337-2
25	211	58.8	74	5	US-10-953-856-57
26	211	58.8	82	5	US-10-953-856-55
27	211	58.8	84	5	US-10-953-856-7

28	210	58.5	67	4	US-10-156-761-11981	Sequence 11981, A
29	209	58.2	70	4	US-10-282-122A-56326	Sequence 56326, A
30	208.5	58.1	66	4	US-10-282-122A-57510	Sequence 57510, A
31	207.5	57.8	66	4	US-10-282-122A-60873	Sequence 60873, A
32	207.5	57.8	66	5	US-10-894-438-4	Sequence 4, Appl1
33	206	57.4	67	4	US-10-156-761-12310	Sequence 12310, A
34	204	56.8	67	4	US-10-156-761-8431	Sequence 8431, Ap
35	202.5	56.4	65	5	US-10-953-856-49	Sequence 49, Appl
36	201.5	56.1	67	5	US-10-953-856-95	Sequence 95, Appl
37	201.5	56.1	67	4	US-10-282-122A-45991	Sequence 45991, A
38	201	56.0	67	4	US-10-156-761-11688	Sequence 11688, A
39	201	56.0	70	4	US-10-282-122A-68525	Sequence 68525, A
40	200.5	55.8	66	5	US-10-953-856-45	Sequence 45, Appl
41	200.5	55.8	67	5	US-10-953-856-65	Sequence 65, Appl
42	200.5	55.8	67	5	US-10-953-856-91	Sequence 91, Appl
43	200.5	55.8	71	5	US-10-953-856-61	Sequence 61, Appl
44	200.5	55.8	79	4	US-10-724-972A-4120	Sequence 4120, Ap
45	200.5	55.8	79	5	US-10-953-856-59	Sequence 59, Appl

ALIGNMENTS

```
RESULT 1
US-09-791-171-147
; Sequence 147, Application US/09791171
; Patent No. US20020094336A1
; GENERAL INFORMATION:
; APPLICANT: ANDERSEN, Peter
; APPLICANT: NIELSEN, Rikke
; APPLICANT: OSTINGER, Thomas
; APPLICANT: RASMUSSEN, Peter Blak
; APPLICANT: ROSENKRANDS, Ida
; APPLICANT: WELDKING, Karin
; APPLICANT: FLORIO, Walter
; TITLE OF INVENTION: NUCLEIC ACIDS FRAGMENTS AND POLYPEPTIDE FRAGMENTS
; TITLE OR INVENTION: DERIVED FROM M. TUBERCULOSIS
; FILE REFERENCE: 670001-2002.1
; CURRENT APPLICATION NUMBER: US/09/791,171
; CURRENT FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 09/050,739
; PRIOR FILING DATE: 1998-03-30
; PRIOR APPLICATION NUMBER: 0376/97
; PRIOR FILING DATE: 1997-04-02
; PRIOR APPLICATION NUMBER: 1277/97
; PRIOR FILING DATE: 1997-11-10
; PRIOR APPLICATION NUMBER: 60/044,624
; PRIOR FILING DATE: 1997-04-18
; PRIOR APPLICATION NUMBER: 60/070,488
; PRIOR FILING DATE: 1998-01-05
; NUMBER OF SEQ ID NOS: 173
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 147
; LENGTH: 67
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-09-791-171-147
Query Match 100.0%; Score 359; DB 3; Length 67;
Best Local Similarity 100.0%; Pred. No. 1.6e-38;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 1 MPOGTVKMFNAEKGFPIADPESGAVFYHTYTGTFGTLEENKVEFEIGHSPKQ 60
DB 1 MPOGTVKMFNAEKGFPIADPESGAVFYHTYTGTFGTLEENKVEFEIGHSPKQ 60
QY 61 ATGVRSI 67
DB 61 ATGVRSI 67

RESULT 2
US-09-804-380-147

```

; Sequence 147, Application US/09804980
; Publication No. US20030147897A1
; GENERAL INFORMATION:
; APPLICANT: Statens Serum Institut
; APPLICANT: Andersen, Peter
; TITLE OF INVENTION: M. Tuberculosis Antigens
; FILE REFERENCE: 670001-2002.4
; CURRENT APPLICATION NUMBER: US/09/804,980
; NUMBER OF SEQ ID NOS: 257
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 147
; LENGTH: 67
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-09-804-980-147

Query Match      100.0%; Score 359; DB 3; Length 67;
Best Local Similarity 100.0%; Pred. No. 1.6e-38;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MPOGTVMFNAEKGFPIAPEDGSADVFWHYTEIOGTGRTLEENQKVFEEIGHSPKGPQ 60
      |||
DB      1 MPOGTVMFNAEKGFPIAPEDGSADVFWHYTEIOGTGRTLEENQKVFEEIGHSPKGPQ 60
      |||

QY      61 ATGVRSL 67
      |||
DB      61 ATGVRSL 67
      |||

RESULT 3
US-10-620-246-147
; Sequence 147, Application US/10620246
; Publication No. US20040115211A1
; GENERAL INFORMATION:
; APPLICANT: ANDERSEN, Peter
; APPLICANT: NIELSEN, Rikke
; APPLICANT: OESTINGEN, Thomas
; APPLICANT: RASMUSSEN, Peter Blak
; APPLICANT: ROSENKRANDS, Ida
; APPLICANT: WELDINGH, Karin
; APPLICANT: PIORIO, Walter
; TITLE OF INVENTION: NUCLEIC ACIDS FRAGMENTS AND POLYPEPTIDE FRAGMENTS
; FILE REFERENCE: 670001-2002.1A
; CURRENT APPLICATION NUMBER: US/10/620,246
; CURRENT FILING DATE: 2003-07-15
; PRIOR APPLICATION NUMBER: 09/050,739
; PRIOR FILING DATE: 1998-03-30
; PRIOR APPLICATION NUMBER: 0376/97
; PRIOR FILING DATE: 1997-04-02
; PRIOR APPLICATION NUMBER: 1277/97
; PRIOR FILING DATE: 1997-11-10
; PRIOR APPLICATION NUMBER: 60/044,624
; PRIOR FILING DATE: 1997-04-18
; PRIOR APPLICATION NUMBER: 60/070,488
; PRIOR FILING DATE: 1998-01-05
; PRIOR APPLICATION NUMBER: 10/138,473
; PRIOR FILING DATE: 2002-05-02
; PRIOR APPLICATION NUMBER: 09/791,171
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 09/415,884
; PRIOR FILING DATE: 1999-10-08
; PRIOR APPLICATION NUMBER: 60/116,673
; PRIOR FILING DATE: 1999-01-21
; PRIOR APPLICATION NUMBER: 1281/98
; PRIOR FILING DATE: 1998-10-08
; NUMBER OF SEQ ID NOS: 173
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 147
; LENGTH: 67
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
```

```

US-10-620-246-147

Query Match      100.0%; Score 359; DB 4; Length 67;
Best Local Similarity 100.0%; Pred. No. 1.6e-38;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MPOGTVMFNAEKGFPIAPEDGSADVFWHYTEIOGTGRTLEENQKVFEEIGHSPKGPQ 60
      |||
DB      1 MPOGTVMFNAEKGFPIAPEDGSADVFWHYTEIOGTGRTLEENQKVFEEIGHSPKGPQ 60
      |||

QY      61 ATGVRSL 67
      |||
DB      61 ATGVRSL 67
      |||

RESULT 4
US-09-738-626-3695
; Sequence 3695, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 3695
; LENGTH: 67
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-3695

Query Match      78.6%; Score 282; DB 3; Length 67;
Best Local Similarity 77.6%; Pred. No. 1.4e-28;
Matches 52; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

QY      1 MPOGTVMFNAEKGFPIAPEDGSADVFWHYTEIOGTGRTLEENQKVFEEIGHSPKGPQ 60
      |||
DB      1 MPOGTVMFNAEKGFPIAPEDGSADVFWHYTEIOGTGRTLEENQKVFEEIGHSPKGPQ 60
      |||

QY      61 ATGVRSL 67
      |||
DB      61 AAGVRSL 67
      |||

RESULT 5
US-10-953-856-51
; Sequence 51, Application US/10953856
; Publication No. US20050097640A1
; GENERAL INFORMATION:
; APPLICANT: Monsanto Technology
; TITLE OF INVENTION: Methods for enhancing abiotic stress tolerance in plants and
; FILE REFERENCE: Docket number (38-21)51768C
; CURRENT APPLICATION NUMBER: US/10/953,856
; CURRENT FILING DATE: 2004-09-29
; NUMBER OF SEQ ID NOS: 95
```

```

; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 51
; LENGTH: 67
; TYPE: PRF
; ORGANISM: Corynebacterium glutamicum
US-10-953-856-51
```

```

Query Match      78.6%; Score 282; DB 5; Length 67;
Best Local Similarity 77.6%; Pred. No. 1.4e-28;
Matches 52; Conservative 5; Mismatches 10; Indels 0; Gaps 0;
```

```

QY      1 MPOGTVMFNAKGFPGFIAPEDGSADVFWYHTTEIQSGFRTLENQVFEFEGHSPKGPQ 60
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      1 MPOGTVMFNAKGFPGFIAPEDGSADVFWYHTTEIQSGFRTLENQVFEFEGHSPKGPQ 60
```

```

QY      61 ATGVRSI 67
      |||:|
Db      61 AQAIVRAL 67
```

```

RESULT 6
US-09-738-626-3845
; Sequence 3845, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738, 626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 3845
; LENGTH: 67
; TYPE: PRF
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-3845
```

```

Query Match      76.6%; Score 275; DB 3; Length 67;
Best Local Similarity 77.6%; Pred. No. 1.2e-27;
Matches 52; Conservative 4; Mismatches 11; Indels 0; Gaps 0;
```

```

QY      1 MPOGTVMFNAKGFPGFIAPEDGSADVFWYHTTEIQSGFRTLENQVFEFEGHSPKGPQ 60
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      1 MPOGTVMFNAKGFPGFIAPEDGSADVFWYHTTEIQSGFRTLENQVFEFEGHSPKGPQ 60
```

```

QY      61 ATGVRSI 67
      |||:|
Db      61 AQAIVRAL 67
```

```

RESULT 7
US-10-282-122A-62113
; Sequence 62113, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
```

```

; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl
; APPLICANT: Zykkind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 62113
; LENGTH: 67
; TYPE: PRF
; ORGANISM: Mycobacterium avium
US-10-282-122A-62113
```

```

Query Match      71.3%; Score 256; DB 4; Length 67;
Best Local Similarity 64.2%; Pred. No. 3.3e-25;
Matches 43; Conservative 14; Mismatches 10; Indels 0; Gaps 0;
```

```

QY      1 MPOGTVMFNAKGFPGFIAPEDGSADVFWYHTTEIQSGFRTLENQVFEFEGHSPKGPQ 60
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      1 MPOGTVMFNAKGFPGFIAPEDGSADVFWYHTTEIQSGFRTLENQVFEFEGHSPKGPQ 60
```

```

QY      61 ATGVRSI 67
      |||:|
Db      61 AQAIVSTV 67
```

```

RESULT 8
US-10-156-761-12354
; Sequence 12354, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
```

```
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 12354
; LENGTH: 67
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-12354
```

```
Query Match          65.5%; Score 235; DB 4; Length 67;
Best Local Similarity 65.7%; Pred. No. 1.7e-22;
Matches 44; Conservative 5; Mismatches 18; Indels 0; Gaps 0;
```

```
QY 1 MPOGTVMFAEKGFPIAPEDGSADVFWHTYBTIOGTGRTLEENQVBEFGHSPKGPQ 60
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 MATGTVMFAEKGFPIAGGGGPDVFWHTYSAINAGFRSLSEENQSVSDVTGQGPQ 60
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 61 ATGVRSL 67
Db 61 AENVTPPL 67
```

```
RESULT 9
US-10-156-761-11466
; Sequence 11466, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; PRIOR FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 11466
; LENGTH: 68
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-11466
```

```
Query Match          64.2%; Score 230.5; DB 4; Length 68;
Best Local Similarity 70.8%; Pred. No. 6.8e-22;
Matches 46; Conservative 5; Mismatches 13; Indels 1; Gaps 1;
```

```
QY 1 MPOGTVMFAEKGFPIAPEDGSADVFWHTYBTIOGTGRTLEENQVBEFGHSPKGPQ 60
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 MATGTVMFAEKGFPIA-VDGDADVFWHTYSAIQMDGYRTLEEGQVDFEISQGXKGPQ 59
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 61 ATGVR 65
Db 60 ADMVR 64
```

```
RESULT 10
US-10-953-856-19
; Sequence 19, Application US/10953856
; Publication No. US20050097640A1
; GENERAL INFORMATION:
; APPLICANT: Monsanto Technology
; TITLE OF INVENTION: Methods for enhancing abiotic stress tolerance in plants and
; FILE REFERENCE: Docket number (38-21)51768C
; CURRENT APPLICATION NUMBER: US/10/953,856
```

```
; CURRENT FILING DATE: 2004-09-29
; NUMBER OF SEQ ID NOS: 95
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 19
; LENGTH: 70
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-953-856-19
```

```
Query Match          62.1%; Score 223; DB 5; Length 70;
Best Local Similarity 65.6%; Pred. No. 6.6e-21;
Matches 42; Conservative 5; Mismatches 17; Indels 0; Gaps 0;
```

```
QY 4 GTTWMFAEKGFPIAPEDGSADVFWHTYBTIOGTGRTLEENQVBEFGHSPKGPQ 63
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 7 GLVWMFAEKGFPIAPEDGSADVFWHTYBTIOGTGRTLEENQVBEFGHSPKGPQ 66
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 64 VRSL 67
Db 67 VVTL 70
```

```
RESULT 11
US-10-156-761-8364
; Sequence 8364, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; PRIOR FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 8364
; LENGTH: 67
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-8364
```

```
Query Match          61.6%; Score 221; DB 4; Length 67;
Best Local Similarity 64.1%; Pred. No. 1.1e-20;
Matches 41; Conservative 4; Mismatches 19; Indels 0; Gaps 0;
```

```
QY 1 MPOGTVMFAEKGFPIAPEDGSADVFWHTYBTIOGTGRTLEENQVBEFGHSPKGPQ 60
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 MATGTVMFAEKGFPIAGGGGPDVFWHTYSAINAGFRSLSEENQSVSDVTGQGPQ 60
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 61 ATGV 64
Db 61 AENV 64
```

```
RESULT 12
US-10-953-856-13
; Sequence 13, Application US/10953856
; Publication No. US20050097640A1
; GENERAL INFORMATION:
; APPLICANT: Monsanto Technology
; TITLE OF INVENTION: Methods for enhancing abiotic stress tolerance in plants and
; FILE REFERENCE: Docket number (38-21)51768C
; CURRENT APPLICATION NUMBER: US/10/953,856
; CURRENT FILING DATE: 2004-09-29
```


GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: April 14, 2006, 18:37:38 ; Search time 2.871 Seconds
(without alignments)
991.565 Million cell updates/sec

Title: US-10-620-246-147
Perfect score: 359
Sequence: 1 MFGSTVKNFNAEKGFPIAD.....VDFEIGHSPKQATGVRSI 67

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 217505 seqs, 42489236 residues

Total number of hits satisfying chosen parameters: 217505

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Published Applications_AA_New:*
1: /SIDS5/ptodata/2/pubpaa/US06_NEW_PUB pep:*
2: /SIDS5/ptodata/2/pubpaa/US06_NEW_PUB pep:*
3: /SIDS5/ptodata/2/pubpaa/US07_NEW_PUB pep:*
4: /SIDS5/ptodata/2/pubpaa/PCT_NEW_PUB pep:*
5: /SIDS5/ptodata/2/pubpaa/US05_NEW_PUB pep:*
6: /SIDS5/ptodata/2/pubpaa/US10_NEW_PUB pep:*
7: /SIDS5/ptodata/2/pubpaa/US11_NEW_PUB pep:*
8: /SIDS5/ptodata/2/pubpaa/US60_NEW_PUB pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	223	62.1	70	6 US-10-506-192-6	Sequence 6, Appl
2	211	58.8	70	6 US-10-506-192-2	Sequence 2, Appl
3	211	58.8	70	7 US-11-077-619-18	Sequence 16, Appl
4	202	56.3	70	6 US-10-506-192-8	Sequence 8, Appl
5	199.5	55.6	67	7 US-11-077-619-20	Sequence 20, Appl
6	199	55.4	71	7 US-11-077-619-22	Sequence 22, Appl
7	199	55.4	72	6 US-10-506-192-4	Sequence 4, Appl
8	198.5	55.3	66	7 US-11-077-619-98	Sequence 98, Appl
9	188	52.4	80	6 US-10-467-657-3284	Sequence 3284, Ap
10	138	38.4	364	7 US-11-169-041-215	Sequence 215, App
11	92	25.6	67	7 US-11-074-176-302	Sequence 302, App
12	91	25.3	147	6 US-10-821-334-1369	Sequence 1369, Ap
13	89	21.9	146	7 US-11-079-463-7806	Sequence 7806, Ap
14	78.5	21.9	826	6 US-10-821-334-1048	Sequence 1048, Ap
15	65	18.1	351	7 US-11-087-099-6490	Sequence 6490, Ap
16	65	18.1	351	7 US-11-087-099-11534	Sequence 11534, A
17	63.5	17.7	562	7 US-11-188-298-5316	Sequence 5316, App
18	63.5	17.7	573	7 US-11-188-298-444	Sequence 444, App
19	61.5	17.7	573	7 US-11-188-298-20628	Sequence 20628, A
20	61.5	17.1	431	6 US-10-467-657-2532	Sequence 2532, Ap
21	61.5	17.1	497	7 US-11-018-868-12	Sequence 12, Appl
22	61	17.0	1475	6 US-10-330-773-182	Sequence 182, App
23	60.5	16.9	576	7 US-11-098-686-10763	Sequence 10763, A
24	60.5	16.9	974	7 US-11-096-568A-33586	Sequence 33586, A
25	60.5	16.9	1032	7 US-11-096-568A-33585	Sequence 33585, A

26	60	16.7	154	7 US-11-188-298-7875	Sequence 7875, Ap
27	59.5	16.6	314	7 US-11-188-298-6359	Sequence 6359, Ap
28	59.5	16.6	314	7 US-11-188-298-10200	Sequence 10200, A
29	59.5	16.6	314	7 US-11-188-298-11910	Sequence 11910, A
30	59.5	16.6	314	7 US-11-188-298-17481	Sequence 17481, A
31	59.5	16.6	461	7 US-11-188-298-4166	Sequence 4166, Ap
32	59.5	16.6	461	7 US-11-188-298-10851	Sequence 10851, A
33	59.5	16.6	461	7 US-11-188-298-13172	Sequence 13172, A
34	59.5	16.6	461	7 US-11-188-298-15364	Sequence 15364, A
35	58.5	16.3	314	7 US-11-188-298-1224	Sequence 1224, Ap
36	58.5	16.3	1000	7 US-11-098-686-10907	Sequence 10907, A
37	58	16.2	191	7 US-11-082-389-26	Sequence 26, Appl
38	58	16.2	235	7 US-11-096-568A-30604	Sequence 30604, A
39	58	16.2	296	7 US-11-096-568A-27880	Sequence 27880, A
40	58	16.2	327	7 US-11-096-568A-27879	Sequence 27879, A
41	58	16.2	368	7 US-11-096-568A-27878	Sequence 27878, A
42	57.5	16.0	314	7 US-11-188-298-5412	Sequence 5412, Ap
43	57.5	16.0	314	7 US-11-188-298-10981	Sequence 10981, A
44	57.5	16.0	314	7 US-11-188-298-16521	Sequence 16521, A
45	57.5	16.0	314	7 US-11-188-298-16630	Sequence 16630, A

ALIGNMENTS

```
RESULT 1
US-10-506-192-6
Sequence 6, Application US/10506192
Publication No. US20050272924A1
GENERAL INFORMATION:
APPLICANT: INOYE, Masayori
APPLICANT: XIA, Bing
APPLICANT: PHADTARE, Sangita
APPLICANT: Qing, Guoliang
APPLICANT: KE, Haijing
TITLE OF INVENTION: COLD SHOCK INDUCIBLE EXPRESSION AND PRODUCTION OF HETEROLOGOUS
FILE REFERENCE: INOYE=2.1 PCT
CURRENT APPLICATION NUMBER: US/10/506,192
PRIOR FILING DATE: 2004-09-01
PRIOR APPLICATION NUMBER: 60/361,069
PRIOR FILING DATE: 2002-03-01
PRIOR APPLICATION NUMBER: 60/402,921
PRIOR FILING DATE: 2002-08-14
NUMBER OF SEQ ID NOS: 50
SOFTWARE: PatentIn version 3.3
SEQ ID NO 6
LENGTH: 70
TYPE: PRT
ORGANISM: Escherichia coli
FEATURE:
OTHER INFORMATION: Synthetic
US-10-506-192-6
Query Match 62.1%; Score 223; DB 6; Length 70;
Best Local Similarity 65.6%; Pred. No. 5e-21;
Matches 42; Conservative 5; Mismatches 17; Indels 0; Gaps 0;
QY 4 GTYKWFNAEKGRFIPDESSADVFTYHTYIGTGRTLEENQKVERIGHSPKQATG 63
DB 7 GLVKNFADKGFIFIPDDSKDVFHFTAIQSNRFTLNNQKVERFSICGGRGPAAN 66
QY 64 VRSL 67
DB 67 VVTL 70
RESULT 2
US-10-506-192-2
Sequence 2, Application US/10506192
Publication No. US20050272924A1
GENERAL INFORMATION:
APPLICANT: INOYE, Masayori
```

QY	64	VRSL	67
Db	67	VTSL	70

US-11-077-619-20

Query Match	55.6%;	Score 199.5;	DB 7;	Length 67;
-------------	--------	--------------	-------	------------

NUMBER OF SEQ ID NOS: 50

APPLICANT: MONACI Elisabetta

```
/ TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
/ FILE REFERENCE:
/ CURRENT APPLICATION NUMBER: US/10/467,657
/ CURRENT FILING DATE: 2003-08-11
/ PRIOR APPLICATION NUMBER: GB-0103424.8
/ PRIOR FILING DATE: 2001-02-12
/ NUMBER OF SEQ ID NOS: 9218
/ SOFTWARE: SeqWIn99, version 1.04
/ SEQ ID NO 3284
/ LENGTH: 80
/ TYPE: PRT
/ ORGANISM: Neisseria gonorrhoeae
US-10-467-657-3284
```

```
Query Match          52.4%; Score 188; DB 6; Length 80;
Best Local Similarity 48.5%; Pred. No. 1.4e-16;
Matches 32; Conservative 13; Mismatches 21; Indels 0; Gaps 0;
```

```
QY 1 MFGTVKMFNAEKGFPGFIAPEDGSADVFWYHTYIQTGFRTLEBNOKVEFEGHSPKGPQ 60
DB 14 MATGIYKMFNDAGFGFPTDEGGEDLFAHFSAINNEGFKTLKGGQVSPDVTGPKGKQ 73
QY 61 ATCVRS 66
DB 74 AANIQA 79
```

```
RESULT 10
US-11-169-041-216
/ Sequence 216, Application US/11169041
/ Publication No. US20060019284A1
/ GENERAL INFORMATION:
```

```
/ APPLICANT: Bristol-Myers Squibb Company
/ TITLE OF INVENTION: IDENTIFICATION OF POLYNUCLEOTIDES FOR PREDICTING ACTIVITY OF
/ TITLE OF INVENTION: COMPOUNDS THAT INTERACT WITH AND/OR MODULATE PROTEIN TYROSINE
/ TITLE OF INVENTION: KINASES AND/OR PROTEIN TYROSINE KINASE PATHWAYS IN LUNG CANCER
/ TITLE OF INVENTION: CELLS
/ FILE REFERENCE: 10001 NP
/ CURRENT APPLICATION NUMBER: US/11/169,041
/ CURRENT FILING DATE: 2005-06-28
/ PRIOR APPLICATION NUMBER: 60/584,405
/ PRIOR FILING DATE: 2004-06-30
/ NUMBER OF SEQ ID NOS: 527
/ SOFTWARE: PatentIn version 3.2
/ SEQ ID NO 216
/ LENGTH: 364
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-11-169-041-216
```

```
Query Match          38.4%; Score 138; DB 7; Length 364;
Best Local Similarity 43.1%; Pred. No. 1.7e-09;
Matches 28; Conservative 9; Mismatches 24; Indels 4; Gaps 1;
```

```
QY 4 GTVKMFNAEKGFPGFIAPEDGSADVFWYHTYIQTG---FRTLEBNOKVEFEGHSPKGP 59
DB 96 GTVKMFNVRNGYGFIRNNDTKEDVFWYHTAIKRNPKFARSVGGDETVFDEVEGRKA 155
QY 60 QATGV 64
DB 156 EATNV 160
```

```
RESULT 11
US-11-074-176-302
/ Sequence 302, Application US/11074176
/ Publication No. US20050250135A1
/ GENERAL INFORMATION:
/ APPLICANT: Kleaenhamer, Todd R.
/ APPLICANT: Russell, William M.
/ APPLICANT: Altermann, Eric
/ APPLICANT: McAvilife, Olivia
/ APPLICANT: Peril, Andrea Azcarate
```

```
/ TITLE OF INVENTION: Nucleic Acid Sequences Encoding
/ TITLE OF INVENTION: Stress-Related Proteins and Uses Therefore
/ FILE REFERENCE: 5051-694
/ CURRENT APPLICATION NUMBER: US/11/074,176
/ CURRENT FILING DATE: 2005-03-07
/ PRIOR APPLICATION NUMBER: 60/551,161
/ PRIOR FILING DATE: 2004-03-08
/ NUMBER OF SEQ ID NOS: 381
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 302
/ LENGTH: 67
/ TYPE: PRT
/ ORGANISM: Lactobacillus acidophilus
US-11-074-176-302
```

```
Query Match          25.6%; Score 92; DB 7; Length 67;
Best Local Similarity 29.7%; Pred. No. 0.00013;
Matches 19; Conservative 12; Mismatches 33; Indels 0; Gaps 0;
```

```
QY 1 MFGTVKMFNAEKGFPGFIAPEDGSADVFWYHTYIQTGFRTLEBNOKVEFEGHSPKGPQ 60
DB 1 MRLGTVKQFPDSSYGFIBDDKHKSIFYFYRAIKEGYKSIRVGDVRYQLAOGKNGIQ 60
QY 61 ATCV 64
DB 61 CVNV 64
```

```
RESULT 12
US-10-821-234-1369
/ Sequence 1369, Application US/10821234
/ Publication No. US20050255114A1
/ GENERAL INFORMATION:
```

```
/ APPLICANT: Labat, Ivan
/ APPLICANT: Stache-Crain, Birgit
/ APPLICANT: Andarmari, Susan
/ APPLICANT: Tang, Y. Tom
/ TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
/ FILE REFERENCE: 821A
/ CURRENT APPLICATION NUMBER: US/10/821,234
/ CURRENT FILING DATE: 2004-04-07
/ PRIOR APPLICATION NUMBER: US 60/462,047
/ PRIOR FILING DATE: 2003-04-07
/ NUMBER OF SEQ ID NOS: 1704
/ SOFTWARE: pc_seq_genes Version 1.0
/ SEQ ID NO 1369
/ LENGTH: 147
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-821-234-1369
```

```
Query Match          25.3%; Score 91; DB 6; Length 147;
Best Local Similarity 44.1%; Pred. No. 0.00045;
Matches 15; Conservative 7; Mismatches 12; Indels 0; Gaps 0;
```

```
QY 3 QGTVKMFNAEKGFPGFIAPEDGSADVFWYHTYIQTG 36
DB 64 KGVCKCFCKSKGKHGFITPADGGPDIPLHISDVEG 97
```

```
RESULT 13
US-11-079-463-7806
/ Sequence 7806, Application US/11079463
/ Publication No. US20060073161A1
/ GENERAL INFORMATION:
/ APPLICANT: Gary L. Breton
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO BACTERIOIDES PC
/ TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
/ FILE REFERENCE: PATH00-03DIV2
/ CURRENT APPLICATION NUMBER: US/11/079,463
/ CURRENT FILING DATE: 2005-03-14
/ PRIOR APPLICATION NUMBER: US 60/128,705
/ PRIOR FILING DATE: 1999-04-09
```

; PRIOR APPLICATION NUMBER: US 09/540,209
; PRIOR FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 10444
; SEQ ID NO 7806
; LENGTH: 146
; TYPE: PRT
; ORGANISM: B.fragilis
US-11-079-463-7806

Query Match 24.8%; Score 89; DB 7; Length 146;
Best Local Similarity 31.1%; Pred. No. 0.00079;
Matches 19; Conservative 10; Mismatches 28; Indels 4; Gaps 1;

QY 4 GTVKWFNAKGFPIAPEDGSADVFWHTYETIQGTGFRITLBNQKVEFEIGHSPKGPQATG 63
DB 85 GRIBYFPAAGYGVNADGGEKFFHISAPA----TIAEGDVTETFEIRGWMGNNAVR 140

QY 64 V 64
DB 141 I 141

RESULT 14
US-10-821-234-1048
; Sequence 1048, Application US/10821234
; Publication No. US20050255114A1
; GENERAL INFORMATION:
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Andatman, Susan
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
; FILE REFERENCE: 821A
; CURRENT APPLICATION NUMBER: US/10/821,234
; CURRENT FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: US 60/462,047
; PRIOR FILING DATE: 2003-04-07
; NUMBER OF SEQ ID NOS: 1704
; SOFTWARE: pt SEQ_genes Version 1.0
; SEQ ID NO 1048
; LENGTH: 826
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-821-234-1048

Query Match 21.9%; Score 78.5; DB 6; Length 826;
Best Local Similarity 37.5%; Pred. No. 0.14;
Matches 21; Conservative 7; Mismatches 27; Indels 1; Gaps 1;

QY 12 EKGGFIAPEDSADVFWHTYETIQGTGFRITLBNQKVEFEIGHSPKGPQATGVRSL 67
DB 222 KEAFPIERGDVVKELTFHISEFKG-DLETLQPGDDVEFTIKDRNGKVAIDVRL 276

RESULT 15
US-11-087-099-6490
; Sequence 6490, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B BP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 6490
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Staphylococcus aureus subsp. aureus Mu50
US-11-087-099-6490

Query Match 18.1%; Score 65; DB 7; Length 351;
Best Local Similarity 39.0%; Pred. No. 2.4;

Matches 16; Conservative 3; Mismatches 20; Indels 2; Gaps 1;
QY 20 PEDGSADVFWHTYETIQGTGFRITLBNQKVEFEI--GHSPK 58
DB 24 PLDDKVKIKVHAGICTDITHYEGHKNVFPVTLGHFPG 64

Search completed: April 14, 2006, 18:42:17
Job time : 3.871 secs

This Page Blank (uspto)